Blending Human and Artificial Intelligence to Support Autistic Children's Social Communication Skills

KAŚKA PORAYSKA-POMSTA, University College London, UCL Institute of Education, United Kingdom
ALYSSA M. ALCORN, University College London, UCL Institute of Education, United Kingdom
KATERINA AVRAMIDES, University College London, UCL Institute of Education, United Kingdom
SANDRA BEALE, Brickbeck College, Department of Psychological Sciences, United Kingdom
SARA BERNARDINI, Department of Computer Science, Royal Holloway, United Kingdom
MARY-ELLEN FOSTER, School of Computing Science, University of Glasgow, United Kingdom
CHRISTOPHER FRAUENBERGER, HCI Group, TU Wien, Austria
JUDITH GOOD, Department of Informatics, University of Sussex, United Kingdom
KAREN GULDBERG, School of Education, University of Birmingham, United Kingdom
WENDY KEAY-BRIGHT, Cardiff School of Art and Design, Cardiff Metropolitan University, United Kingdom
LILA KOSSYVAKI, School of Education, University of Birmingham, United Kingdom
OLIVER LEMON, School of Mathematical and Computer Sciences Heriot-Watt University, United Kingdom
MARILENA MADEMTZI, Yale University, Child Study Center, USA
RACHEL MENZIES, University of Dundee School of Science and Engineering, United Kindgom
HELEN PAIN, School of Informatics, University of Edinburgh , United Kingdom
GNANATHUSHARAN RAJENDRAN, Department of Psychology, Heriot-Watt University, United Kingdom
ANNALU WALLER, University of Dundee, School of Science and Engineering, United Kingdom
SAM WASS, University of East London, United Kingdom
TIM J. SMITH, Department of Psychological Sciences, Birkbeck, University of London, United Kingdom

Authors' addresses: Kaśka Porayska-Pomsta, University College London, UCL Institute of Education, 23-27 Emerald Street, London, WC1N 3QS, United Kingdom, K.Porayska-Pomsta@ucl.ac.uk; Alyssa M. Alcorn, University College London, UCL Institute of Education, United Kingdom, a.alcorn@ucl.ac.uk; Katerina Avramides, University College London, UCL Institute of Education, United Kingdom; Sandra Beale, Brickbeck College, Department of Psychological Sciences, United Kingdom; Sara Bernardini, Department of Computer Science, Royal Holloway, United Kingdom, Sara.Bernardini@rhul.ac.uk; Mary-Ellen Foster, School of Computing Science, University of Glasgow, United Kingdom, MaryEllen.Foster@glasgow.ac.uk; Christopher Frauenberger, HCI Group, TU Wien, Austria, christopher.frauenberger@tuwien.ac.at; Judith Good, Department of Informatics, University of Sussex, United Kingdom, J.Good@sussex.ac.uk; Karen Guldberg, School of Education, University of Birmingham, United Kingdom, K.K.Guldberg@bham.ac.uk; Wendy Keay-Bright, Cardiff School of Art and Design, Cardiff Metropolitan University, United Kingdom, wkbright@cardiffmet.ac.uk; Lila Kossyvaki, School of Education, University of Birmingham, United Kingdom; Oliver Lemon, School of Mathematical and Computer Sciences Heriot-Watt University, Edinburgh, United Kingdom, o.lemon@hw.ac.uk; Marilena Mademtzi, Yale University, Child Study Center, USA, marilena.mademtzi@yale.edu; Rachel Menzies, University of Dundee School of Science and Engineering, United Kindgom, r.menzies@dundee.ac.uk; Helen Pain, School of Informatics, University of Edinburgh , United Kingdom, helen@inf.ed.ac.uk; Gnanathusharan Rajendran, Department of Psychology, Heriot-Watt University, Edinburgh, United Kingdom, t.rajendran@hw.ac.uk; Annalu Waller, University of Dundee, School of Science and Engineering, United Kingdom, a.waller@dundee.ac.uk; Sam Wass, University of East London, United Kingdom, s.v.wass@uel.ac.uk; Tim J. Smith, Department of Psychological Sciences, Birkbeck, University of London, United Kingdom, tj.smith@bbk.ac.uk.

This paper examines the educational efficacy of a learning environment in which children diagnosed with Autism Spectrum Conditions (ASC) engage in social interactions with an artificially intelligent (AI) virtual agent and where a human practitioner acts in support of the interactions. A multi-site intervention study in schools across the UK was conducted with 29 children with ASC and learning difficulties, aged 4-14 years old. For reasons related to data completeness and amount of exposure to the AI environment, data for 15 children was included in the analysis. The analysis revealed a significant increase in the proportion of social responses made by ASC children to human practitioners. The number of initiations made to human practitioners and to the virtual agent by the ASC children also increased numerically over the course of the sessions. However, due to large individual differences within the ASC group, this did not reach significance. Although no evidence of transfer to the real-world post-test was shown, anecdotal evidence of classroom transfer was reported. The work presented in this paper offers an important contribution to the growing body of research in the context of AI technology design and use for autism intervention in real school contexts. Specifically, the work highlights key methodological challenges and opportunities in this area by leveraging interdisciplinary insights in a way that (i) bridges between educational interventions and intelligent technology design practices, (ii) considers the design of technology as well as the design of its use (context and procedures) on par with one another, and (iii) includes design contributions from different stakeholders, including children with and without ASC diagnosis, educational practitioners and researchers.

CCS Concepts: • **Human-centred computing** → HCI design and evaluation methods; user studies; • **Education** → Interactive learning environments; • **User characteristics** → children;

Additional Key Words and Phrases: Autism, artificially intelligent agent, social communication, intelligent learning environments, neurodiversity

ACM Reference Format:
Kaśka Porayska-Pomsta, Alyssa M. Alcorn, Katerina Avramides, Sandra Beale, Sara Bernardini, Mary-Ellen Foster, Christopher Frauenberger, Judith Good, Karen Guldberg, Wendy Keay-Bright, Lila Kossyvaki, Oliver Lemon, Marilena Mademtzi, Rachel Menzies, Helen Pain, Gnanathusharan Rajendran, Annalu Waller, Sam Wass, and Tim J. Smith. 2010. Blending Human and Artificial Intelligence to Support Autistic Children's Social Communication Skills. *ACM Comput. Entertain.* 9, 4, Article 39 (March 2010), 38 pages. https://doi.org/0000001.0000001

1 INTRODUCTION

Autism is a spectrum of neuro-developmental disorders that affects the way in which a person communicates with and relates to other people, as well as how they make sense of the world around them [NAS 2016]. The main areas of difficulty in Autism Spectrum Conditions (ASC) are: (i) *social communication and interaction*, involving problems with both verbal and non-verbal language, e.g. difficulty with initiating or responding to bids for interaction, or with taking turns in conversations and (ii) *restricted, repetitive patterns of behaviour, interests, or activities*, for example, difficulties with adapting to novel environments or coping with unexpected change (see [APA 2013] for a comprehensive list of potential difficulties). The degree to which such difficulties are present and their exact nature varies between individuals and possibly across cultural and situational contexts. This heterogeneity necessitates individualised and adaptive support regimes that are sensitive to the individuals' real-life routines and that are not confined to laboratory or clinical intervention settings.

A wide range of clinical and non-clinical interventions exists that aim to facilitate the learning and development of social communication skills. Over the past decade, technology-enhanced methods have attracted increasing attention in both the autism and educational communities for their potential to impact both research and practice [Goodwin

2008; Parsons et al. 2015; Parsons and Mitchell 2017]. Software interventions have targeted language skills, e.g. [Anwar et al. 2011; Bosseler and Massaro 2003; Massaro 2006; Rahman et al. 2011], affective skills, e.g. [Abirached et al. 2009; Beaumont and Sofronoff 2008; Finkelstein et al. 2009; Hopkins et al. 2011; Schuller et al. 2013], and social interaction skills, e.g. [Barakova et al. 2007; Battocchi et al. 2013; Dautenhahn and Werry 2004; Kandalaft et al. 2013; Kozima et al. 2009]. The latter include approaches to fostering social initiation in peer-to-peer collaborative contexts [Malinverni et al. 2014; Mora-Guiard et al. 2016; Tartaro and Cassell 2008], attentional control [Bartoli et al. 2013], and imaginary/symbolic play [Herrera et al. 2008, 2012] through exploratory full body interaction environments. The existing interventions examine the design and use of both advanced Artificial Intelligence (AI), e.g. as in Barakova et al.'s [Barakova et al. 2007] social robotics work, as well as technologically shallower interactive interface designs, e.g. as in Herrera et al.'s [Herrera et al. 2008, 2012] approach to full-body interactions.

The proliferation of different technologies and their applications for autism is being increasingly accompanied by a body of evidence suggesting that technology, in the broad sense of the word and in tightly controlled environments, may provide effective support for this target group [Cobb 2007; Parsons et al. 2007; Parsons and Mitchell 2017]. This is especially true in relation to *within technology use* improvements, e.g. [Wass and Porayska-Pomsta 2014], with some indication of the potential for generalisation to real-world situations [Bosseler and Massaro 2003; Fletcher-Watson 2014; Golan et al. 2010; Grynszpan et al. 2014; Hourcade et al. 2013; Pennington 2010; Ramdoss et al. 2011a,b; Tartaro and Cassell 2008]. Coupled with this emerging evidence is a growing demand from diverse groups, including parents, educators and individuals with ASC, for the adoption of an inclusive support model which views autism as a difference in abilities rather than as a spectrum of deficits. In such an inclusive model, first, any intervention should be adapted to individuals' strengths rather than their deficits, with the therapeutic regimes being personalised to their individual needs. Second, appropriate adjustments should be made not only to the individuals' physical, but also transactional environments to facilitate and scaffold their engagement in social and interpersonal interactions [Prizant et al. 2003]. The focus on strengths rather than deficits, and the importance of adaptable and adaptive environments to the success of interventions, has been long highlighted as key to best educational practice [Biesta 2007; Dewey 1998]. This view also aligns with the aspirations of Artificial Intelligence in Education research, where moment-by-moment adaptation to the idiosyncratic needs and actions of learners is of central interest, e.g. [Woolf 2008]. This, together with AI's long-term key ambition to emulate human behaviours in socially credible ways, provides a strong motivation for examining the educational and interactional potential of AI technologies in this context.

Given that initiations and responses which are either absent or inappropriate are seen as key areas of difficulty in ASC [APA 2013], this paper focuses on these two behaviours. Children with ASC tend to initiate all types of communication infrequently compared with *typically developing* (TD) or developmentally delayed peers [Mundy et al. 2003], and respond to partners in restricted or self-serving ways. A key aim of the work presented here was to create an environment in which the child was encouraged and motivated to produce *spontaneous* communication behaviours, i.e. to engage in "communication in the absence of [a] defined antecedent" [Chiang and Carter 2008]. Such behaviours are treasured by parents and practitioners, because they occur infrequently and when they do, they indicate that the child acknowledges them as social agents.

This paper presents the design and evaluation of a technology called ECHOES, which opportunistically blends human and AI support for autistic children, aged 4-14 years old[1], in their exploration of social interaction skills. Such blending aims to deliver rich and flexible transactional support to children as advocated in the autism best practices.

[1]The ECHOES environment was originally designed for children aged 4-7 years old. However, as explained in Section 4, several of the study participants were chronologically older, but developmentally within or below the target age range

Here, AI serves as a stepping stone for the social interactions with the humans, whereby human practitioners provide on-demand support when a child is willing and able to interact and communicate with them, or where a combination of technology and human intelligence is necessary to cater adequately for the interaction needs of the specific child.

The contributions of the work presented are four-fold. *First*, the study provides a detailed example of how human-computer interaction can be orchestrated in a way that blends human and artificial intelligence for the benefit of each individual child. This is a particularly timely contribution, given the emerging evidence-driven trend in AI in Education of adopting carefully blended AI-human approaches for supporting learning [Baker 2016]. *Second*, through the examination of the efficacy of the ECHOES approach, the study spotlights the potential of technology as a trigger and a catalyst for meaningful social communication between ASC children and adults in non-clinical contexts such as schools. *Third*, in contrast with much research in technology-enhanced approaches to autism intervention, the work extends existing research by targeting children at the lower end of the autism spectrum conditions. *Fourth*, the research presented highlights important methodological challenges related to the ways in which any improvements in autistic children's social communication may be measured meaningfully in a way that (i) informs and innovates front-line practices, (ii) points to how educational interventions for this population may be designed and delivered specifically with the help of AI technologies, and (iii) informs how the definition of optimal outcome for children with ASC may need to be framed to allow for the design and delivery of more inclusive support regimes than are currently routinely available (for further discussion see Sections 4 and 7.4).

The rest of the paper is organised as follows. Section 2 describes the ECHOES environment, explaining its pedagogical underpinnings. Section 3 describes the design and functionality of the ECHOES AI agent. Section 4 introduces the study design along with the research questions addressed. Section 5 presents the ECHOES data annotation scheme, while Section 6 provides the results of the study. Section 7 discusses the results by exploring their broader implications to technology-enhanced interventions for autism, and by outlining key methodological considerations related to enabling inclusive interventions and technological designs in this context. The concluding remarks are given in Section 8 together with examples of work that has already emerged from the ECHOES project, illustrating the rich interdisciplinary basis that the work presented here offers for future work more generally. Supplementary materials related to the ECHOES system architecture and detailed examples of the ECHOES learning activities are provided in the Appendix (A).

2 THE DESIGN OF THE ECHOES ENVIRONMENT

ECHOES is a single user technology-enhanced learning environment that utilises an artificially intelligent virtual character[2] , called Andy, as a social partner for children with ASC and their typically developing peers, to help them learn and/or improve social communication skills (see Fig.1; for details of Andy's intelligence, see e.g. [Bernardini and Porayska-Pomsta 2013]). The design of ECHOES was optimised to:

(1) Encourage and support behavioural change, through:
   (a) a child-centred approach;
   (b) exploration and play opportunities;
   (c) potential interaction with social partners (i.e. virtual or human partners), providing opportunities for the child to initiate and respond to social communication;

[2]The terms 'artificial agent' and 'virtual character' will be used interchangeably in this paper.

Fig. 1. A child playing with the interactive cloud through the ECHOES multi-touch display

(2) Deliver learning activities based on existing evidence of best practice in autism – in this case the Social Communication, Emotional Regulation and Transactional Support (SCERTS) model [Prizant et al. 2006] (see Section 2.1 for details);
(3) Be informed by input from stakeholders, including children with and without autism, practitioners, and experts in developing technology for children with ASC;
(4) Be suitable for use by young children with ASC, who may also have learning difficulties[3];
(5) Be suitable for use in school environments.

In ECHOES, children can both explore different situations involving the artificial agent and rehearse them repeatedly. The interaction between the child and the environment is facilitated through a large 42-inch multi-touch screen, which allows children to manipulate different interactive objects on the screen. The interactive objects provide the opportunities for shared attention and interaction between the agent and the child, and between the child and the human practitioner. The scale of the screen allows children to move freely in front of it. The touch interface caters for children's varying motor control capabilities, as well as for young children's low literacy. To reduce the complexity and potential difficulties of executing touch gestures such as those with a timing element (e.g. double tap or long touch) or those requiring multiple touch points (pinch), the input gestures are restricted to the simplest and least ambiguous possibilities (touch, drag, swipe up/down). From the outset, ECHOES was developed for use in schools and specialist units by children aged 4-7 years old, or of corresponding developmental age (in the studies presented, the chronological age was sometimes as old as 14 years).

[3]This is often referred to as 'intellectual disability' outside of the UK.

2.1 Pedagogical Underpinnings of ECHOES

In designing ECHOES, we drew on best practices in autism intervention [Prizant et al. 2006] and consulted with stakeholders throughout, including teachers and other practitioners, children with and without autism, researchers specialising in autism, and experts in technology design for children with ASC and other special needs. In doing so, we adopted a participatory design approach such that stakeholders could engage with the idea of ECHOES in a personally meaningful way, and express their views and ideas about the ways in which the technology might be shaped to accommodate their individual needs.

An early study [Alcorn et al. 2011] alerted us to the potential of ECHOES as an object for social communication: the presence of the researcher, and the children's spontaneous interaction with them, led to the explicit inclusion of the human partner in the interaction with the ECHOES environment, in addition to the AI agent. This blending of human and AI interaction provided two potential social partners for children while they interacted with ECHOES, offering them a richer, more robust and potentially more immediately transferrable experience than might be possible with only the AI agent as the social partner. The design requirements for ECHOES were based on the outputs of fourteen participatory design workshops. Formative evaluation studies, involving children with and without a diagnosis of autism, were carried out using increasingly sophisticated prototypes of ECHOES. In total, over eighty children and more than thirty adults contributed to the ECHOES design process. For details of the individual workshops and their outcomes see [Bernardini et al. 2014; Frauenberger et al. 2013, 2010, 2011; Porayska-Pomsta et al. 2011].

The design of ECHOES was informed by and validated against the SCERTS model [Prizant et al. 2003] – a comprehensive approach to autism assessment and intervention, combining clinical as well as educational approaches. SCERTS addresses the core challenges related to ASC, and aims to support children in developing a number of key skills across three dimensions that are crucial to social interaction: (1) Social Communication (SC): spontaneous and functional communication, emotional expression, and secure and trusting relationships with others; (2) Emotional Regulation (ER): the ability to maintain a well-regulated emotional state to cope with everyday stress and to be available for learning and interacting; and (3) Transactional Support (TS): the development and implementation of support to help caregivers to respond to the child's needs and interests, to modify and adapt the environment, and to provide tools to enhance learning.

SCERTS breaks down these three domains into a number of further components and, for each one, provides a detailed description of the objectives to be achieved, the intervention strategies available to practitioners and parents, and the criteria for assessing the child's current skills and needs. In designing ECHOES and the learning activities therein, we built explicitly on this operationalisation of social communication. We also used the SCERTS assessment protocol as the basis for developing a bespoke annotation scheme, which we used to code the video data of the children's interactions with ECHOES. The SCERTS framework and the ECHOES' annotation scheme developed thereupon are presented in detail in Section 5.

2.2 The ECHOES Learning Activities

Children's interactions with ECHOES are structured around twelve learning activities which focus on social communication and, in particular, on: **(1)** *joint attention*: a child's ability to coordinate and share attention by looking towards people or shifting gaze between people and objects, share emotions by using facial expressions, express intentions, engage in turn-taking, and participate in reciprocal social interactions by initiating/responding to bids for interaction; and **(2)** *symbolic use*: a child's understanding of meaning expressed through conventional gestures, words, and sentences and their ability to use nonverbal means and vocalisations to share intentions. The learning activities correspond directly to the intervention goals specified in the SCERTS framework.

One of SCERTS' principles is that in order to support joint attention, any activities performed with a child need to share "an obvious unifying theme" [Prizant et al. 2003]. Therefore, all ECHOES activities are set in a "magic" garden inhabited by Andy[4] – the AI agent. The ECHOES garden is magical, because it contains interactive objects that behave in unusual ways, sometimes transforming into other objects when touched by the agent or the child. For example, tapping a flower turns it into a floating bubble or a bouncy ball, depending on the type of touch-gesture used. The design of the environment, the objects and their behaviours, was informed by participatory design work conducted with children with ASC, typically developing (TD) children, parents and practitioners, as described is subsection 2.1. The idea behind the unusual/magical behaviours of the ECHOES objects was to create opportunities for sustained attentional interest and exploration, and for joint activities between the child, the agent and/or the human social partner.

In order to offer children a structured experience when interacting with ECHOES, whilst also leaving room for the children's own expression and discovery, we designed two sets of activities: (i) *closed-task* activities, with a clear sequence of steps and a predetermined end-point; and (ii) *exploratory* activities, with no predetermined end-point and whose main objectives are social reciprocity, turn taking and mutual enjoyment. For example, sorting a set of balls by colour or collecting a predefined number of flowers are examples of closed-task activities which end once all of the objects are either sorted or collected. In contrast, taking turns with the agent to shake a virtual cloud that produces rain and causes flowers to grow, or throwing virtual bouncy balls through the cloud to make them change colour, represent exploratory activities which can go on for as long as the child wishes or is deemed productive by the accompanying practitioner (see Table 1 for a list of all ECHOES activities and the possible interactions between the child, the agent and the human respectively; see also the Appendix A.2 for detailed examples).

Most of the learning activities in ECHOES are performed by Andy and the child in cooperation, with Andy assuming a more or less prominent role according to a particular activity's learning objective and the needs of the individual child. For example, if the goal is learning-by-imitation, Andy will adopt a leading role and will demonstrate different behaviours to the child. If the goal is engaging-in-reciprocal-interaction, Andy will give the child an opportunity to initiate a bid for interaction by waiting before initiating the interaction. There are two activities in ECHOES, which do not involve Andy: (i) bubble popping and (ii) free exploration of the magic garden. One or both of these activities are always used at the beginning of each session (see Fig.2 for a possible sequence of activities over the course of an ECHOES session). The child has full control over the interactions in these activities and our early pilot studies showed that children found these exploratory activities particularly rewarding [Alcorn et al. 2011].

As described subsection 2.1, a human partner is included in the interaction with the ECHOES environment. The child has opportunities to interact with both the AI agent and the human partner. It is important to note that the interaction with the agent was a pre-requisite for the child being able to complete the specific tasks. In contrast, interaction with the human was not required, and took the form of further encouragement and praise, demonstration of actions by the practitioners, and practitioners' readiness to respond to children's bids for interaction, e.g. when a child invited the practitioner to take a turn within an activity or pointed to something of interest on the screen. As the classroom is the intended environment for ECHOES, the human partner will be referred to as the practitioner, although in the study described in Section 4, this role was frequently adopted by a researcher.

[4]Henceforth, we will refer to Andy, as Andy, the virtual character or the agent interchangeably.

Table 1. Inventory of the ECHOES activities. Bubble popping and Magic Garden free-play activities take place at the beginning of all sessions. Other activities are used based on each child's preferences (e.g. if a child likes a particular activity, it is likely that they will play with it more frequently) and whether or not the child has played with particular activities before, as the aim is for all children to play with all ECHOES activities.

ACTIVITY	TYPE	OBJECTS	INTERACTION WITH AGENT	INTERACTION WITH HUMAN
Bubble popping	Exploratory	Bubbles	No	Practitioner either demonstrates that bubbles can be popped or reacts to the child's request to pop bubbles
Magic Garden	Exploratory	Magic cloud can rains; flowers; flower basket; pots; bouncy balls	No	Practitioner demonstrates how objects can be manipulated. Responds to bids for interaction from the child
Ball Sorting	Closed-Task *Reciprocal interaction*	Balls and coloured boxes	Andy: - demonstrates how to sort balls; – instructs the child; – praises the child; – signals end of the task	Practitioner further demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity
Tickling	Exploratory	None	Andy responds to being tickled by bending over and laughing	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity
Explore with Andy	Exploratory *Learning by imitation*	Magic cloud that can produce rain; flowers; flower basket; pots; bouncy balls	Andy: – demonstrates the actions on objects – takes turns with the child to play with the objects – praises the child	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity
Rainy Cloud	Exploratory *Learning by imitation*	Magic cloud; flowers that grow when the cloud is shaken	Andy: – demonstrates how shaking cloud produces rain and this causes the flowers to grow – takes turns with the child to shake the cloud – praises the child	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity
Flower growing	Closed-Task *Reciprocal interaction*	Magic cloud; flowers that grow when the cloud is shaken	Andy: – demonstrates the actions on objects – takes turns with the child – praises the child – Signals end of the task	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity
Pick flowers	Exploratory *Learning by imitation*	Flowers, basket	Andy: – demonstrates the action – takes turns with the child to fill basket with flowers – asks for particular flowers to be picked – praises the child	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity
Stack pots	Closed-Task *Reciprocal interaction*	Pots	Andy: – demonstrates the action – takes turns with the child – asks for a particular pot to be stacked – praises the child – Signals end of the task	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity
Fill basket with flowers	Closed-Task *Reciprocal interaction*	Flowers, basket	Andy: – demonstrates the action – takes turns with the child – asks for a particular flower to be put in the basket next – praises the child for completing action – Signals end of the task	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity
Turn flower into a ball	Exploratory *Learning by imitation*	Flowers; bouncy balls Flower centres change into bouncy balls when flicked	Andy: – demonstrates the actions – takes turns with the child to turn flowers into bouncy balls – praises the child	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity; helps the child
Throw balls through the cloud	Closed-Task *Reciprocal interaction*	Cloud, bouncy balls; Change colour when they go through the cloud	Andy: -demonstrates the action -takes turns with the child -asks for a particular ball to be thrown through the cloud -praises the child -Signals end of the task	Practitioner demonstrates the activity if needed; responds to child's requests to share the activity; encourages the child to engage in the activity

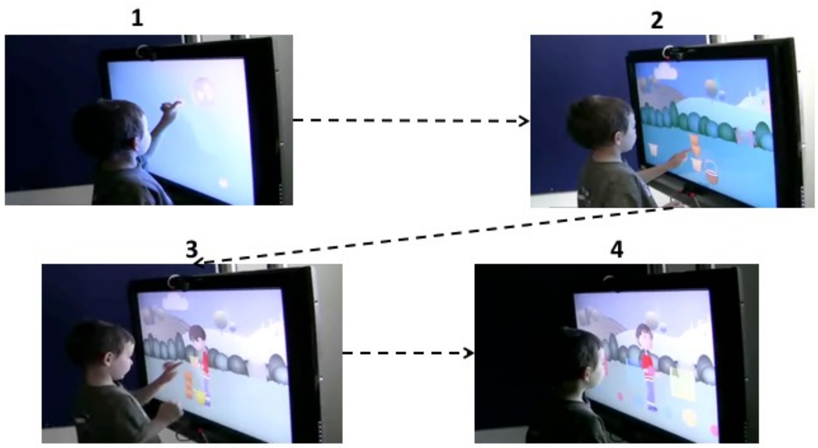

Fig. 2. A possible sequence of activities in one session. (1) bubble popping; (2) magic garden free play, without Andy; (3) stacking pots with Andy; (4) ball sorting. The bubbles activity was always first, frequently followed by exploration without Andy. For children who are unable to concentrate over a long session, an activity with Andy directly followed bubbles.

3 THE ECHOES AI AGENT AS A SOCIAL PARTNER

Following the principles of autism and educational best practices, children need a responsive social partner. Such a partner needs to be able to model the behaviours targeted, to provide a socially credible real-time interaction, while also being able to engage children in structured learning. In other words, children need a social partner that is both (i) pro-active, i.e. has an ability to exhibit goal-directed behaviour by deliberating about, and actively trying to accomplish, its goals and by taking the initiative, and (ii) reactive, i.e. has the ability to perceive changes in the environment and react to them in a timely manner. Furthermore, such a partner needs to possess social ability to coordinate its actions with the actions of others – in our case, the child. Pro-activeness is important to maintaining the child's attentional focus and to fostering motivation. Reactivity is fundamental to adapting to children's changing needs and their cognitive and affective states. Social ability is crucial to maximising the child's sense of self-efficacy in communicating with the social partner. These requirements are fully in line with intelligent agent theory by Wooldridge and Jennings [Wooldridge and Jennings 1995]. The theory, which continues to provide the basis for the design of intelligent agents, offers strong motivation for developing an autonomous planning-based agent within ECHOES that is able to act as a believable and educationally viable social partner to children. Such investment is further supported by existing evidence of the generalisable therapeutic and educational potential of virtual agents, e.g. [Bosseler and Massaro 2003; Grynszpan et al. 2008; Parsons and Cobb 2011; Strickland et al. 2007; Tartaro and Cassell 2008] and by ongoing successful research into applications of planning architectures as the basis for furnishing synthetic, socially enabled characters with autonomous and believable behaviours [Anderson et al. 2013; Aylett et al. 2009; Cavazza et al. 2002; Kenny et al. 2007].

Therefore, the key component of the ECHOES environment is an autonomous planning-based agent which drives the decision-making of the ECHOES AI agent. The agent utilises the FAtiMA planning architecture [Dias and Paiva 2005] that is based on: (i) AI planning techniques [Russell and Norvig 2003], (ii) an emotional model derived from the OCC cognitive theory of emotions [Ortony et al. 1988], and (iii) the appraisal theory of emotions [Smith and Lazarus

1990]. It is a domain-independent agent architecture that provides the ECHOES agent with deliberative and reactive capabilities along with some basic socio-emotional competences. The two main mechanisms controlling a FAtiMA agent are appraisal and coping. Both mechanisms work at the reactive level, which affects the short-term horizon of the agent's behaviours, e.g. agent's timely reactions to a child's momentary actions, and at the deliberative level, which relates to the agent's long-term goal-oriented behaviours, e.g. scaffolding a child to complete a learning activity. Within each activity the agent's actions are controlled by the ECHOES Intelligent Engine and are based on the sequence of real-time updates produced by the Multimodal Fusion Engine. Further details of the rationale for the ECHOES software architecture and individual components are given in [Foster et al. 2010] and [Bernardini and Porayska-Pomsta 2013].

The agent's actions are either concrete demonstrations of social-communication skills targeted in ECHOES or actions performed to encourage the child to try and to practice them. Specifically, we define joint attention and symbolic use in terms of three component skills: (i) responding to bids for interaction; (ii) initiating bids for interaction; (iii) engaging in turn taking. Andy, the agent, is able to demonstrate these skills in three different ways:

(1) Verbally – by using simple language or keywords (e.g. *"My/Your turn!"* for turn-taking);
(2) Non-verbally – through gaze and gestures such as pointing at or touching an object, or using one of a small number of Makaton gestures[5] (e.g. thumbs up for *"Well done"*);
(3) By combining verbal and non-verbal behaviours.

Andy is able to make requests, greet the child by name, comment on events happening in the magic garden and apply exploratory actions to the interactive objects. This variety of behaviours is intended to make the interaction dynamic enough to keep the child engaged and to foster generalisation of skills practiced, while retaining a degree of predictability that is essential to supporting the child's attentional focus and sense of safety. Andy always provides the child with positive feedback, especially if the child responds to his bids for interaction correctly. If the child does not perform the expected action, Andy first waits for the child to do things at their own pace and then intervenes by demonstrating the action. To support the child?s interaction within specific activities, Andy always explains a new activity to the child using simple language and precise instructions (e.g., *"Let's pick ten flowers"*). The importance of providing the children with positive feedback was emphasised by teachers who contributed to the design of Andy's behaviours as key to reducing children's anxiety in social interactions and helping them experience a sense of self-efficacy.

ECHOES also incorporates a practitioner graphical user interface (henceforth referred to as the GUI) through which the practitioner can control the choice and duration of the learning activities. The GUI aims to cater for the substantial diversity in individual children's needs and preferences. As well as enhancing the ECHOES' real-time interpretation of the child's needs, for example the need to repeat an activity that a child might have found particularly enjoyable or to make Andy repeat a specific action if the practitioner deemed it beneficial for a particular child, the GUI also provides a tool for creating an environment in which no opportunity is lost to engage children in social interaction and communication. The GUI is accessed through a separate screen, not visible to the child, so as not to interfere with the flow of the interaction.

4 STUDY DESIGN

We conducted a multi-site intervention study in real school contexts to examine the potential impact of the ECHOES environment use on social communication skills in children with ASC. Specifically, we focused on children's ability to

[5]Makaton is a language programme using signs and symbols to help people with language and learning difficulties to communicate. https://www.makaton.org/

initiate or respond to bids for interaction by others [APA 2013]. In doing so, we examined the *efficacy* of the ECHOES environment, i.e. its potential to have a desired effect, as opposed to its *effectiveness*, i.e. definitive proof that it led to the desired outcome, in this case to the acquisition of social interaction skills (see e.g. [Rao et al. 2008] for an elaboration of the definitions of efficacy vs. effectiveness). This distinction means that we have not been able to control for all factors that may potentially have confounded our findings. However, we believe that the methods we employed to create, deploy and evaluate the ECHOES approach are in line with observations and postulates for autism intervention and education in-the-wild, e.g. [Parsons et al. 2013].

The study used a within-subjects design, looking at changes in initiations/responses across condition (ECHOES vs. non-ECHOES) and over time (the beginning, middle and end phases of interaction with ECHOES). As we were particularly interested in how initiation/response patterns changed over time for children with ASC, a matched control group was neither appropriate, nor feasible, given the profiles of our participants. Many of the children with ASC who participated in the ECHOES studies were characterised as having learning difficulties, confirmed by pre-test assessment (see Section 4.1), and most had been assessed as developmentally delayed. In terms of their language ability, many were at the social (SP) or language partner (LP) stage, and the others were at the conversational partner (CP) stage, as defined within SCERTS. Owing to their profiles, there can be no functionally appropriate comparison group for these children and hence, the traditional 'group-matching' method is not appropriate for direct comparison, particularly when it is based on IQ (see [Dennis et al. 2009] and [Rao et al. 2015] for further discussions of why this is problematic).

Nevertheless, the so-called "lower-functioning end of the spectrum" is often under-researched, overlooked and perhaps even under-estimated, e.g. [Mottron 2004]. It is therefore important to ensure that individuals considered to be "low functioning" are represented in research, to avoid bias in the literature. Furthermore, it may be that this group has the greatest potential to benefit from environments such as ECHOES.

Similar to Whalen and Schreibman [Whalen and Schreibman 2003], who included a group of typically developing children in their study in order to "identify 'normal' levels of social behaviours", we also included a group of typically developing children in the study. They were not considered as a control group, but allowed us to obtain a broad measure of "typical" social interactions (i.e. initiations/responses) across the various study conditions, in effect providing a non-clinical point of reference.

Our specific research questions for the study were as follows:

(1) Do ASC children show an increased response to bids for social interaction while using the ECHOES environment?
  (a) Does this pattern differ between the virtual agent and human practitioner?
(2) Do ASC children show an increase in the number of initiations for social interaction made whilst using the ECHOES environment?
  (a) Does this pattern differ between the virtual agent and human practitioner?
(3) Do any increases in response/initiations transfer to other contexts?

4.1 Participants

Five schools and specialist units for children with ASC were identified and invited to participate in the study. The study design was pilot tested in a further school over 2-3 sessions with four children with ASC. Following revisions to the study design and to the system based on the pilot, ECHOES was deployed at four sites (five schools) across the UK: three special schools and two mainstream schools. One of the mainstream schools had a unit dedicated to working with children with ASC and other learning difficulties. The second shared a site with a school for children with severe, complex and enduring additional support needs, including autism. A group of typically developing (TD) children was recruited from each of the two mainstream schools. Fig. 3 provides an overview of initial and final numbers of child participants, including information on the tests administered.

Twenty-nine children with a previous ASC diagnosis were exposed to ECHOES across the four sites, and their interactions were video-recorded. These children had previously received a diagnosis of autism via qualified authorities and professionals. Children recruited from the three special schools had also been assessed as having learning difficulties. Although it would not have been either appropriate or ethical to re-diagnose the children solely for the purpose of this project, we used the Social Communication Questionnaire (SCQ) and British Picture Vocabulary Scales (BPVS) to gather further evidence about the children's specific difficulties.

Children who had less than 45 minutes of interaction with ECHOES, or who did not take part in a minimum of three ECHOES sessions, were excluded from the analysis, reducing the ASC participant numbers from 29 to 19. The main reason for attrition was illness and other absence. Additionally, a small number of children chose not to continue the use of ECHOES at different points, therefore their data was too incomplete to provide a meaningful basis for analysis. The children who did not want to play with ECHOES were not distinguished from the remaining children by any diagnostic traits except for potential over-sensitivity to the sensory features of ECHOES, such as garden sounds. Other factors may have played a role, for example, children being away from their typical classroom environment, lighting in the room and time of day and week – afternoons and end of the week tended to be the most difficult owing to children's tiredness.

Also excluded were children who did not participate in both pre- and post 'table-top' assessments, further reducing this to 15 children with ASC (see Fig. 3). In addition, the TD children in one of the two mainstream schools worked in pairs (at the school?s request) so their data was not comparable with individual use, thus they were also excluded from the analysis. As a result, a subset of fifteen ASC children was included in the analysis presented here.

The ASC group ($N = 15$) had an average age of 8.54 years (range 4 to 14 years) and included one girl. Confirmation of the ASC diagnosis was provided by the Social Communication Questionnaire (SCQ), which was completed by a caregiver. The average SCQ score was 23.4 ($SD = 4.64$), i.e. higher than the cut-off score of 15 (see [Barnard-Brak et al. 2016] for a discussion about cut-off scores for the SCQ regarding sensitivity and specificity). Verbal language ability was assessed using the 2nd Edition of the British Picture Vocabulary Scale (BPVS). The mean raw BPVS score was 36.62, giving a mean age equivalent of 3.99 years ($SD = 0.97$), significantly lower than the group's chronological age. It should also be noted that seven children had BPVS scores that were too low to provide age equivalence. The group of 15 children with ASC included 5 younger children (aged 7 to 8 years) and 5 older children (aged 13 to 14 years) with learning difficulties in addition to the diagnosis of autism. 3 further children were assessed, using BPVS, as having a raw age equivalent more than 1 year below their chronological age. Participant information is provided in Table 2.

Twelve typically developing children were recruited to provide the reference group. As noted above, six were then excluded from the analysis because they had been working in pairs. The average age of the TD group ($N = 6$) was 5.45 years and included three girls. The TD children did not complete the SCQ. Average BPVS raw score was 63.83 and age equivalence was 5.89 years. Both the age and BPVS scores differed significantly between TD and ASC groups ($p < .001$).

The study design was approved by the ethics board of the site managing the evaluation, and the approval was further reviewed and accepted by all of the participating sites. Ethics approval covered all participatory design, formative evaluation and other studies, in addition to the multiple site study presented in this paper. In all cases information sheets and consent forms were provided to potential participants, parents/guardians, practitioners and teachers. Language appropriate consent forms and information sheets were provided for all children, and read aloud to them if necessary.

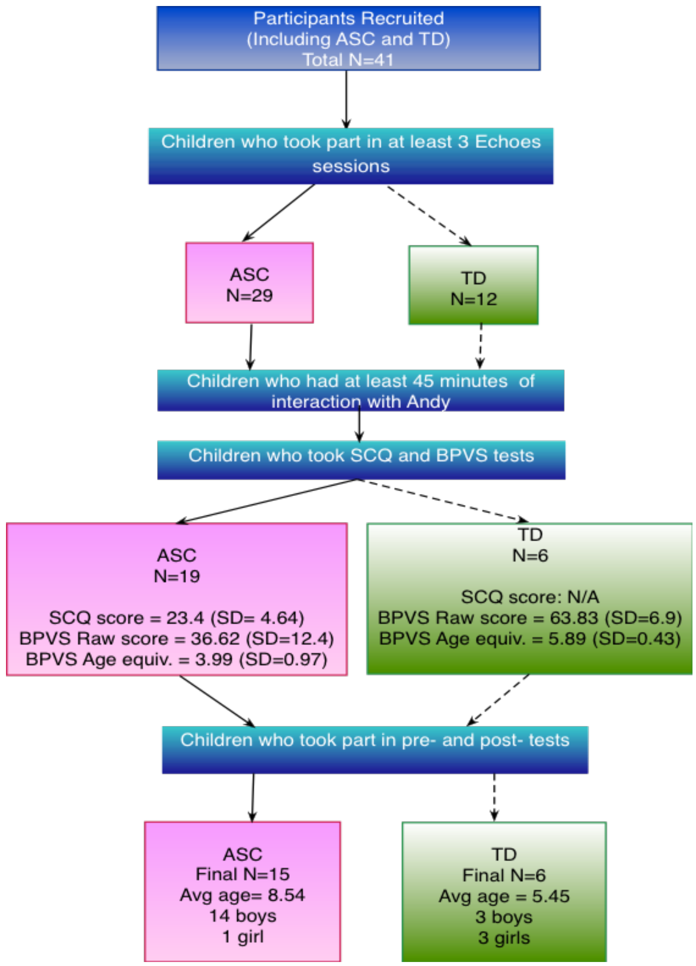

Fig. 3. A summary of participants' selection, numbers and the corresponding tests administered.

4.2 Materials and Procedure

Given that the study was to be conducted concurrently on multiple sites, researchers were trained in the application of the study procedure through workshops and reference to the ECHOES "Manual for Researchers", which was specially developed by the evaluation team to ensure consistency across all sites. The training workshops and manual covered all aspects of the study procedures, including training in using the BPVS, checklists for preparation for working in schools and control of the ECHOES environment. The manual also contained detailed plans for the study, including guidance on the placement of cameras, selection and order of activities, and interacting with and supporting the participants. In addition, a detailed specification was developed in relation to data collection and storage.

Prior to the children's interaction with ECHOES, the researchers assigned to each school spent several sessions observing the children with whom they would be working in the classroom settings. The researchers participated in classroom activities, allowing the children to become familiar with them. ECHOES was installed in quiet, dedicated

Table 2. Participant data was anonymised; school type is Learning Difficulties (LD) or Mainstream (MS); for Language Level see above. BPVS: 2 were not completed due to behavioural difficulties (NC); 3 were assessed as below the 3-year age cut-off for this version. SCQ: for 2 children data was provided by the school as >15 cut-off.

PCode	School Type	SCERTS Level (SP, LP, CP)	Age: years, months	Gender	SCQ score	BPVS age equiv.
RD11	LD/ASC	LP-CP	7.6	M	23	3.3
AM12	LD/ASC	CP	8.2	M	23	5.0
OM16	LD/ASC	LP	7.8	M	25	3.2
GJ26	LD/ASC	LP	13.7	M	18	NC
SD27	LD/ASC	LP	13.1	M	17	4.2
CM29	LD/ASC	SP-LP	13.8	M	32	3.2
RD30	LD/ASC	CP	13.8	M	28	NC
ALH31	LD/ASC	CP	12.8	F	26	5.11
HB51	MS/ASC	SP	4.10	M	27	<3
HK53	MS/ASC	LP	5	M	23	3.2
KL54	MS/ASC	CP	5.11	M	14	5.10
DSS55	MS/ASC	CP	5.9	M	20	4.3
EW56	MS/ASC	CP	5.5	M	23	4.3
MG78	LD/ASC	LP	6.8	M	>15	<3
MD80	LD/ASC	LP	7.4	M	>15	<3
SM57	MS/TD	–	5.5	F	–	6.0
GW58	MS/TD	–	5.2	F	–	6.6
MD59	MS/TD	–	5.6	F	–	5.0
NO60	MS/TD	–	5.6	M	–	6.7
TW61	MS/TD	–	5.4	M	–	6.9
HD62	MS/TD	–	5.4	M	–	6.9

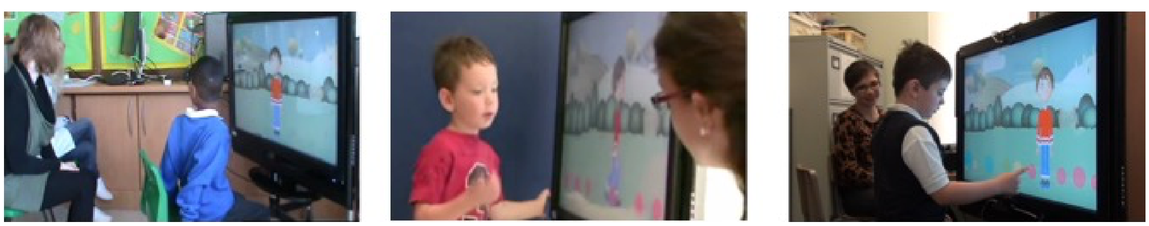

Fig. 4. Children using ECHOES with the researcher at three different evaluation sites.

spaces in each school. Individual children interacted with the environment whilst their interaction was monitored and structured by the human partner/practitioner (Fig. 4). Depending on their preferences, children sat on a chair or stood in front of the screen while the practitioner (and classroom assistant when necessary and available) sat to the side of the screen out of the child's immediate line of sight. Practitioners were able to control the various aspects of the ECHOES environment if needed through the GUI designed for this purpose (see also Section 3). At three of the five sites, the researchers acted as practitioners. At the other two sites, teachers and teaching assistants were trained to take on this role, and worked in conjunction with the researchers.

To assess each child's initial social communication skills, their behaviour was video-taped during (i) free-play at school (e.g. in the playground), (ii) usual classroom activities, (iii) a structured group turn-taking exercise in the classroom, and (iv) a structured one-on-one table-top turn-taking activity. These also served as familiarisation activities, allowing the children to get to know the researchers for the purpose of reducing children's anxiety in preparation for the sessions with ECHOES. Behaviours observed in these videos were coded and quantified using the SCERTS-based ECHOES annotation scheme described in Section 5.

The structured table-top turn-taking activity involved the child playing with two toys (a bubble gun and a remote-controlled robot) on a table-top, and a human practitioner (Fig. 5). The design of these activities was informed by the behavioural assessments used in diagnostic tools such as ADOS [Lord et al. 2012]. The practitioner was instructed to take turns with the child in controlling the toy, using joint attention and pointing, in order to direct the child?s attention to objects on the table-top and to respond to any bids for interaction made by the child. They were also instructed to provide opportunities for the child to initiate, e.g. by remaining quiet for short periods.

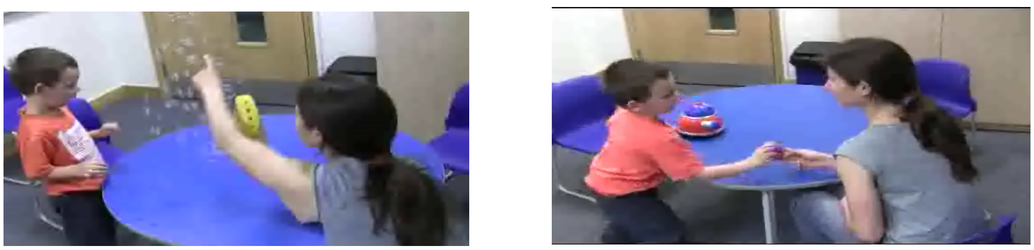

Fig. 5. The table-top activity used to assess social communication skill before (pre) and after (post) ECHOES use. Left = bubbles activity. Right = robot activity.

After the initial table-top pre-test, each child was given the opportunity to play with the ECHOES environment for periods of 10-20 minutes, several times a week over a six-week period. The structure of each session was decided collaboratively by the child and the practitioner, with the practitioner suggesting which learning activities to engage with and when to move between them. All sessions began with the bubbles activity and progressed to the magic garden free-play. However, the time spent on each activity was decided by the practitioner or upon child's request, depending on how engaged the child appeared to be. As the number of activities increased in each subsequent session, less time was spent on these two initial activities. The overview of the entire procedure employed in ECHOES is shown in Fig. 6.

The complexity of the activities increased with each new session, and the ECHOES agent was gradually introduced to act as a social partner to the child. The practitioners were instructed to progress to new activities when they judged the child had grasped the current activity, had become bored or upon child's request. Learning activities included in the evaluation were: *Bubble popping, Magic garden free-play, Ball sorting, Tickling Andy, Explore with Andy, Rainy Cloud, Flower Growing, Pick flowers, Stack pots, Turn flower into a ball, Turn flowers into bubbles* and *Throw balls through the magic cloud.* The AI agent was used in most of the learning activities, but the extent to which he was critical to the goal of an activity varied. For instance, during the early exploratory activity Andy would be present and children could tickle him or hand him objects but he would not initiate interaction. Later activities, such as sorting coloured balls into boxes, required turn-taking: Andy would point to request specific balls which the child could hand to him.

If the child became distressed, distracted or disengaged from ECHOES during a session, the session would be terminated and the child would return to their class. In the small number of cases where this happened, and impeded the

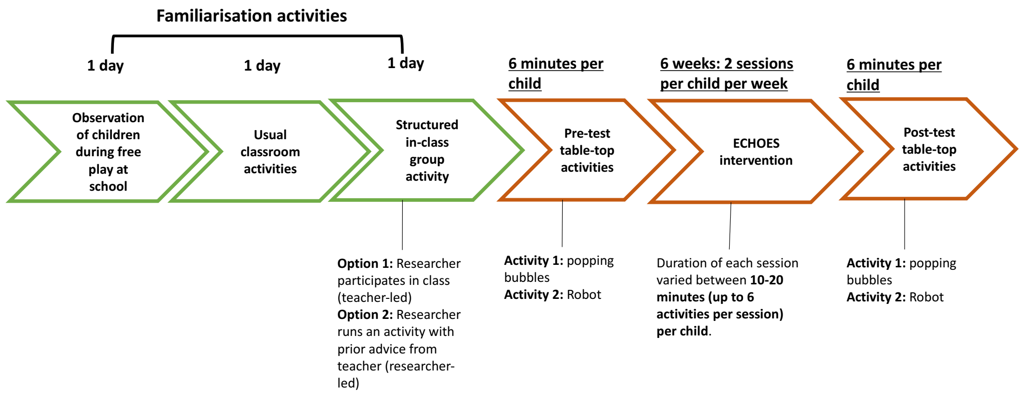

Fig. 6. ECHOES study procedure, detailing the pre-/post- testing and main intervention along with specific timings for each.

child's use of ECHOES, the child's data was removed from the final analysis sample, as already reported in subsection 4.1.

All ECHOES sessions were video-taped. At the end of the six weeks of children using ECHOES, the pre-test activities were repeated. For conciseness, we report only the results of the table-top activities.

5 DATA CODING SCHEME: ADAPTED SCERTS ASSESSMENT PROTOCOL (SAP)

The coding system used to assess the video data collected during the summative evaluation is an adapted version of the SCERTS Assessment Process (SAP) [Prizant et al. 2006]. The original SAP is a curriculum-based assessment tool for gauging individual children's capacity to use certain skills and to engage in tasks across meaningful, everyday contexts. The emphasis is on understanding the functional role of an individual child's behaviours and communicative acts, rather than solely identifying the deficits. Assessment of a child's level of social and communicative competence is based on detailed behavioural criteria derived from the developmental and autism literature, rather than by reference to group norms (see overview in [Prizant et al. 2006]: Volume 1, Chapter 7).

Despite its orientation toward rapid, on-going assessment in practice, the SAP provided an invaluable starting point for the data coding in ECHOES, owing to (i) the wide range of socio-communicative behavioural criteria included in the SAP, and (ii) the fact that many high-level behaviours (e.g. requesting) have already been decomposed into smaller, individual criteria (e.g. request food, request help, request comfort). Moreover, these criteria are context-independent, and are therefore relevant to a wide range of interactions in which there may be no prior list of "relevant" or "possible" child behaviours. This is very different from the narrower, context-dependent data coding schemes for assessments such as ADOS [Lord et al. 2012], or the Early Social Communication Scales (ESCS) [Mundy et al. 2003], which dictate the application of pre-specified, standardised materials to very structured one-on-one interactions between the child and an adult "tester", whose task is to elicit specific behaviours from the child. The SAP's range and flexibility is appropriate as a starting point for the varied ECHOES evaluation contexts (classroom data, structured table-top activities, and interaction with the virtual environment) because it allows us to account for both the use of technology-enhanced learning and the real-world school context of the study. Whilst other coding schemes, such as [Bauminger 2002] and [Hauck et al. 1995], are designed to be used in natural environments, their coding schemes do not provide the hierarchical levels of categories, or the operationalisation of the lower levels of detail, that SAP does.

In the modified SAP for ECHOES coding scheme (SAP-E), we have largely kept the spirit and the behavioural criteria of the original SAP, while making some changes in order to render it more useful as a research tool in the context of ECHOES. In particular, the SAP-E uses a modified subset of the behavioural codes of the original SCERTS framework, and adds several new codes that capture additional information specific to ECHOES. The behavioural codes are applied incrementally to allow us to capture information about a child's behaviours at multiple levels of detail. For example, at a high level, a given behaviour may be coded as a verbal response to the virtual character, but can also be coded as constituting a greeting, and as being an instance of exact echolalia (i.e. imitation of another person's speech). The current paper reports the results related to the higher-level codes, rather than the specific results for these more detailed codes.

The main changes between SAP and SAP-E can be summarised as follows:

(1) Shift to counting instances of behaviours instead of estimating frequency of use;
(2) Limit coding to ECHOES-relevant socio-communicative behaviours;
(3) Provide more detailed, stringent definitions of child initiation;
(4) Add codes that define social partners' initiations and responses to the child;
(5) Add codes that define "missed opportunities" for the child to respond to social partners.

While these changes may seem substantial, we still consider the SAP-E to be an adapted and extended version of the SAP, rather than a completely new coding scheme. This is because the SAP-E's focus remains on investigating communicative competence in a flexible, unscripted interaction that, in addition to the child, may involve multiple social partners and objects at various points. Furthermore, most of SAP-E's behavioural codes follow the SAP verbatim, or have very minor alterations (i.e. to specify that they apply to the virtual character, or to remove parts of the code clearly irrelevant to ECHOES). The main changes introduced through SAP-E relate to bridging the gap between the formative, practice-focused uses of the SAP, and the definitional precision and contextual information required to turn it into a tool able to produce research-quality data about child communication in a technology-enhanced environment.

5.1 Applying the coding scheme to ECHOES data

Video from each participant included in the analysis was sampled at the beginning, middle and end of the ECHOES intervention period, excluding child rest breaks, technical malfunctions, and footage of other activities not specifically related to ECHOES. Each sample analysed comprised 15 minutes of child-ECHOES interaction using activities in which the agent was present, as well as the transitions between such activities. Samples were selected as follows: beginning (beg) samples started in the session where the virtual agent was first introduced, ending (end) samples started with the virtual agent's exit from the final session with ECHOES (with the 15 minutes of qualifying footage counted back from that point), while the middle (mid) sample was drawn from the child's middle session with the virtual agent.

Each video was first coded by a researcher trained in the SAP-E. Training included extensive practice on ECHOES pilot data, iteratively discussing and amending the coding until it was close to the previously coded and agreed 'master' version for one training video. Ten first-coders (selected from the authors and associated students) were used in total, all trained to be able to accurately reproduce the SAP-E coding of the training video before progressing. In coding the final evaluation data, coders were blind to the child?s diagnostic category (ASC or TD) and the video's phase (e.g. pre-test or post-test table-top; beg, mid or end of the ECHOES intervention), though differences in context such as table-top versus ECHOES were self-evident. Codes were applied using the ELAN Linguistic Annotator . ELAN allows multiple, overlapping codes to be attached to sections of video through a mixture of free-text entry and menus of pre-defined labels, facilitating a high degree of flexibility in identifying and labelling child-partner interactions. After first coding was completed for all videos, all annotations were fully moderated by one of two second coders in order to improve confidence in the final analysis. Whenever moderators disagreed with first codes and could not resolve the conflict easily, the two moderators conferred and reached agreement. Full moderation was used instead of partial second-coding for various reasons: (1) with ten independent coders, any calculation of inter-rater reliability (IRR) would be difficult to derive meaningfully; (2) by having two moderators view all annotations and reach agreement we were aiming for a high level of consistency in the application of the coding scheme across all videos rather than simply confidence in the amount of variation in the codes; (3) the substantial variation in ability of the children documented in the videos combined with the number of coders, meant that after experiments with IRR during piloting of the annotation scheme it became clear that a high-level overview of all children was required when applying the coding scheme. Final annotations were exported as tab-delimited text for further analysis.

5.2 Behavioural codes discussed in this paper

For the purpose of the analysis presented in this paper, and in line with our hypotheses, we focus on the three high-level categories of socio-communicative behaviours that are less frequently used by children with ASC:

(1) Child's response to bids for interaction from a social partner, including:
   (a) Verbal and non-verbal responses to partner's verbal and non-verbal initiations;
   (b) Following the reference of a partner's pointing
(2) Child's initiation of bids for interaction to a social partner, including:
   (a) Use of pointing to direct a partner's attention;
   (b) Requesting (objects, help/actions, activities);
   (c) Protesting undesired actions,
   (d) Other verbal and non-verbal bids not otherwise specified, including commenting on objects or events.
(3) Child's social behaviour towards partners, in particular:
   (a) Using gaze for social referencing (i.e. looking towards a partner for information) and social sharing (i.e. initiating joint attention through a combination of gaze and gesture to convey enjoyment and interest);
   (b) Monitoring the attentional focus of the partner in an on-going activity;
   (c) Securing the attention of the partner;
   (d) Greeting the partner;
   (e) Facilitating continuation of turn taking.

In SAP-E, each behaviour is listed as a high-level category description that encompasses multiple specific codes, and is supported by specific examples of child behaviours. In many cases, a behaviour could feasibly be classified as fulfilling more than one behavioural objective, and hence be coded under all categories that capture the relevant information about that behaviour. The current analysis also incorporates information on social partners' initiations and responses to the child, as well as missed opportunities for children to respond to social partners' bids for interaction (discussed in Section 5). Finally, as an additional and more qualitative source of information, the analysis draws on concrete, i.e. verbatim examples of children's speech directed to the virtual agent throughout their ECHOES sessions. For example, in Fig. 7, the child observes Andy trying to put a blue ball in the yellow box. He leans forward and touches the blue box and tells Andy "Right here!" then, with his other hand, touches the blue box twice and shouts: "That one!". In this example, we see the child initiating a bid for interaction to the AI social partner, and using pointing to direct a partner's attention.

Fig. 7. Left: A child watches Andy make a sorting mistake. Right: The child contact points to the blue box and tells Andy "Right here!".

6 RESULTS

The analysis presented is based on the pre-ECHOES table-top videos, three 15-minute ECHOES sessions involving the AI agent (beginning, middle, and end of the intervention period), and the post-ECHOES table-top sessions. Our analysis focuses on children's initiations of and responses to bids for interaction to social partners, including (i) the AI agent and (ii) the human practitioner. In order to be coded, behaviours must be relevant to the ECHOES system in some way (e.g. the child initiating to tell the researcher about a field trip, say, would not be coded under the SAP-E).

6.1 Frequency of children's responses to a social partner

Given that the number of initiations made by the partner (both the HUMAN practitioner and the AGENT) may vary substantially across sessions and across children, the successful responses made by the child will be expressed as a proportion: ranging from 1 (all initiations responded to) to 0 (no initiations responded to). Given that lesser responsiveness to initiations by a social partner is a diagnostic trait of ASC (e.g. as measured by ADOS), it was predicted that our ASC children would exhibit lower proportions of responses both during the pre-ECHOES table-top session and throughout ECHOES, as compared to their typically developing peers. Of interest here is whether these proportions change, and how they vary across the HUMAN practitioner – a socially complex, but more reactive and reliable social partner, and the AGENT – a simpler, more predictable partner, but with limited interactivity compared to the human.

The mean proportion of ASC children's responses to the practitioner's bids for interaction during the table-top pre-test was $0.66 (SD = 0.17)$, i.e. significantly less than the reference proportion for TD children (overall $mean = 0.87, SD = 0.17, t(21) = -3.412, p < 0.01$) . This confirms that our ASC sample began the study with significant impairment in social responsiveness compared to an unmatched TD sample.

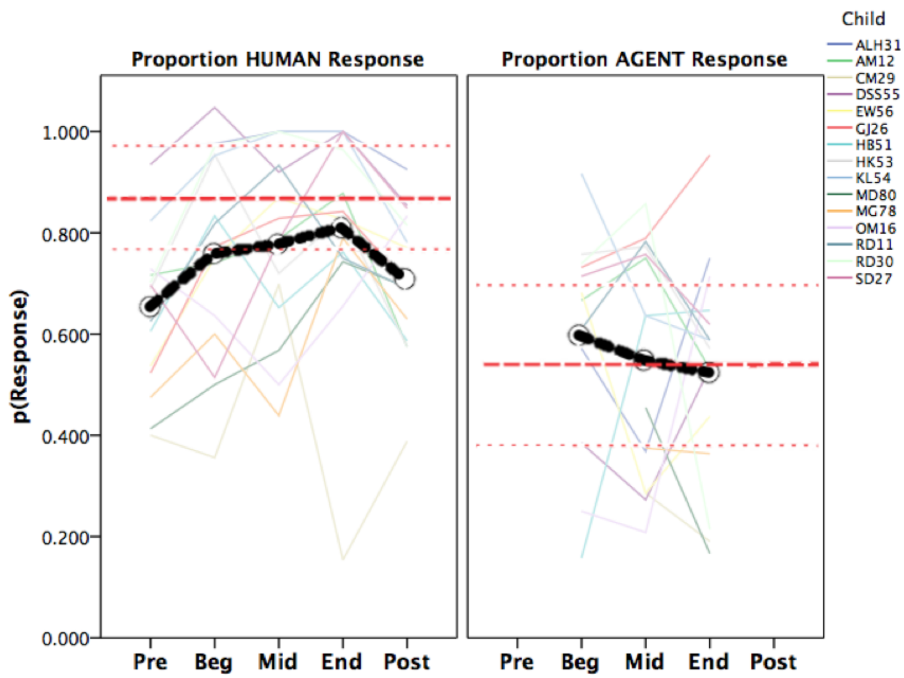

Fig. 8. Mean proportion of child responses to partner initiations (black dashed line), split by the HUMAN practitioner (left panel) or Andy the AGENT (right panel) across the pre- table-top, beginning, middle and end ECHOES sessions and the post- table-top session. Reference mean proportion of TD children's responsiveness to initiations is represented as the red dashed line (the average proportion of responses over all five time periods) with 95% confidence intervals for the TD group represented as red dotted lines. Background coloured lines represent individual participant values at each session (anonymised IDs tally with Table 2).

What is clear from the mean proportion of responses displayed in Fig. 8 is the increase in responses during the use of ECHOES. A repeated-measures ANOVA performed on the proportion of responses to HUMAN initiations across the 5 phases (pre and post table-top, and beginning, middle, end of ECHOES sessions) showed a significant main effect of phase ($F(4, 56) = 4.167, p < 0.01$), which can be attributed to the end time point showing a greater proportion of response ($mean = 0.813, SD = 0.214$) compared to the pre-test ($mean = 0.656, SD = 0.171; t(14) = -3.72, p < 0.01$). This increase also brings the ASC responsiveness to a level comparable to that of the TD group. However, this increase in responsiveness does not transfer outside of ECHOES to the table-top activity, as shown by the post-test ($mean = 0.711$, $SD = 0.142$), the results of which are not significantly different to the pre-test, $t(14) = -1.637, p = 0.124, n.s.$).

The key difference seen in Fig. 8 is the low proportion of child responses to AGENT initiations (Fig. 8, right panel; $mean = 0.59, SD = 0.22$) compared to the responses to the HUMAN practitioner (Fig. 8, left panel; 0.8). This difference is significant within the ECHOES sessions ($F(1, 25) = 9,588, p < 0.01$) and may be due to issues with Andy's temporal contingency on the child's interactions. Specifically, while the planning architecture used to underpin the agent's behaviours gave the agent decision-making capabilities, and enabled greater responsiveness to the child, the agent's reactions were at times slightly delayed owing to a current plan needing to complete execution before it could react to the child's next action. This was particularly an issue with children who did not observe turn-taking and did not wait for Andy to finish what he was doing. The slight decrease in response proportion over time does not reach significance

($F < 1$). Note that the TD reference responsiveness to Andy (Fig. 8, right panel, red dashed line) was similarly low, suggesting that the ASC children did not find Andy any more difficult as a social partner than the TD children.

6.2 Frequency of children's initiations to a social partner

Initiating can be considered more difficult than responding to a partner, because it relies crucially on the initiator's motivation to communicate with another person as well as their ability to communicate without prompting. In contrast, the form and the function of the expected type of response will be directly or indirectly prompted by the partner's initiation, giving the child clues not only that a response is expected, but also often what form of response is required.

The videos were coded for initiation behaviours directed to both the HUMAN practitioner and the AGENT, and considered across verbal, non-verbal and combined initiations. Initiation behaviour is more difficult to standardise as (i) we do not know how many initiations should be expected per child and per session, and (ii) the different learning activities and level of ability of each child may substantially impact the frequency of their initiations.

The results of the analysis focusing on the frequency of initiations to HUMAN, illustrated in Fig. 9 (left panel) by a black dashed dotted line, reveal that the ASC children make more initiations before and during ECHOES use than would normally be expected given the profile of this group. The mean frequency of HUMAN initiations for the TD reference group is $3.1 (SD = 2.9)$ per fifteen-minute block, represented as a red dashed line in Fig. 9 (95% confidence intervals denoted by red dotted lines). By comparison, the frequency of initiations for the ASC group is higher than the TD reference across all five phases (independent-samples t-test reveal a significant difference at all phases, $p < 0.05$). This may not be as surprising as it first seems. The TD reference group are not IQ matched to the ASC group and as the ECHOES environment was designed for users with significantly less ability, the TD group found the learning activities easy and did not need to seek the help from, or make comments to the practitioner. Furthermore, in the pre- and post-table-top session, the TD children waited until they were instructed to do something by the practitioner, whereas the ASC children were often making off-task requests for interaction which, although not strictly pre-/post-test task related, could be seen as positive, given ECHOES' aim of increasing ASC children's bids for interaction.

The results of the analysis focusing on the frequency of initiations to HUMAN, illustrated in Fig. 9 (left panel) by a black dashed dotted line, reveal that the ASC children make more initiations before and during ECHOES use than would normally be expected given the profile of this group. The mean frequency of HUMAN initiations for the TD reference group is $3.1 (SD = 2.9)$ per fifteen-minute block, represented as a red dashed line in Fig. 9 (95% confidence intervals denoted by red dotted lines). By comparison, the frequency of initiations for the ASC group is higher than the TD reference across all five phases (independent-samples t-test reveal a significant difference at all phases, $p < 0.05$). This may not be as surprising as it first seems. The TD reference group are not IQ matched to the ASC group and as the ECHOES environment was designed for users with significantly less ability, the TD group found the learning activities easy and did not need to seek the help from, or make comments to the practitioner. Furthermore, in the pre- and post-table-top session, the TD children waited until they were instructed to do something by the practitioner, whereas the ASC children were often making off-task requests for interaction which, although not strictly pre-/post-test task related, could be seen as positive, given ECHOES' aim of increasing ASC children's bids for interaction.

A repeated-measures ANOVA of initiation frequency does not show a significant effect of phase pre, beg, mid, end, post; $F(4, 56) = 1,341, p = 0.266$) even though there is a numerical increase in initiations from pre ECHOES session ($mean = 10.27, SD = 8.1$) to end ECHOES session ($mean = 17.9, SD = 25.9$). This increase does not reach significance due to large variance across children ($t(16) = -1.361, p = 0.192$). This, in turn, highlights (i) the heterogeneity within

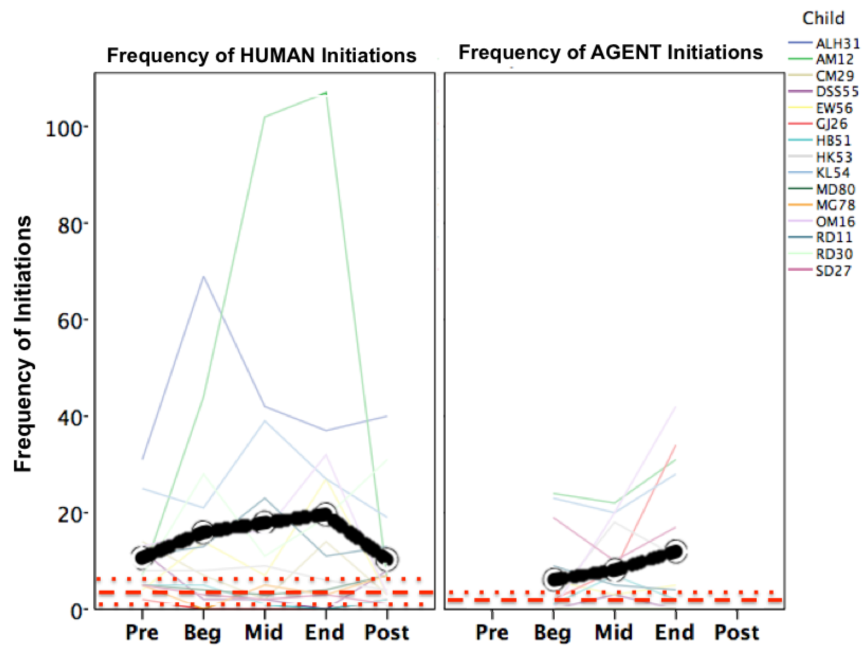

Fig. 9. Mean frequency of initiations made by child (black dashed line) to the HUMAN practitioner (left panel) or to Andy the AGENT (right panel) across the pre- table-top, beginning, middle and end ECHOES sessions and the post- table-top session. Reference frequency of TD children initiations is represented as the red dashed line (the average proportion of responses over all five time periods) with 95% confidence intervals for the TD group represented as red dotted lines. Background coloured lines represent individual participant values at each session (anonymised IDs tally with Table 2).

our ASC group (see the variable frequency of initiations across individual participants represented as the different colour lines in Fig. 9 left panel) and (ii) the fact that some children may have benefited from ECHOES more than others.

In terms of children's initiations to the AGENT, we observe numerically fewer initiations than towards the HUMAN practitioner. However, this difference does not reach significance (F(1,28)=2.072, p=0.161, n.s.). Similar to the initiations to the HUMAN partner, the initiations to Andy also seem to be characterised by an increase over the course of the ECHOES sessions (Fig. 9, right panel, black dashed line). ASC children show an increase from $4.89(SD = 8.05)$ in the first session to $9.6(SD = 13.7)$ in the final ECHOES session, however, this difference does not reach significance due to the large variance across individuals ($t(18) = -1.719, p = .103, n.s.$). For some ASC children, their comfort and interest in Andy may be increasing even though his interactivity is limited, e.g. the individual participant coloured lines above the black dashed group mean line in Fig. 9, right panel. An increase in initiations towards the AGENT should be expected as a natural consequence of progressing through the ECHOES learning activities and moving on to the later activities that require turn-taking with Andy.

7 DISCUSSION

7.1 Frequency of initiations and responses to social partners

Our results show that as some of the children with ASC progress through the ECHOES environment (from beginning, middle and end), they seem to show corresponding numerical increases in initiations of joint-attentional behaviours to both the virtual character and the human practitioner. Arguably, this change may be considered clinically, if not statistically significant at the group level. This is in contrast to the pre- and post- non-ECHOES 'table top' measures of the children's initiations of joint attention. What is of particular note is that initiations have been argued to be the most difficult of all joint attentional behaviours for children with ASC [APA 2013]. Here, the ECHOES environment seems to engender an increase in these behaviours in some ASC children compared to outside the ECHOES environment and, moreover, this change increases as the child spends more time with ECHOES.

With respect to responding to bids for interaction, the pattern is slightly more nuanced in that the children with ASC showed an increasing tendency to respond to their human partner over the course of the ECHOES sessions, but an overall lower level of responsiveness to the AI agent. One explanation for this may be that a child's response to the human partner could lead to continued reciprocal interaction, whereas a response to Andy could not, since Andy was not aware that the child had responded unless the response related to, and was expressed through, the touch actions recognisable by ECHOES. Another possible explanation might be that actually blending human and AI interaction allows children to ease into the situation, making them more open to engaging with the human over time, with the AI agent then transforming from the central interaction vehicle to an object around which the interaction and communication can be conducted with the human partner. Although other researchers have also shown that technology can enhance co-located human interaction, e.g. [Farr et al. 2010; Holt and Yuill 2017], it is worth noting that, in each case, these study designs explicitly included more than one individual interacting with the technology at the same time, either as pairs (in the case of [Holt and Yuill 2017]) or groups of individuals (in the case of [Farr et al. 2010]). In contrast, ECHOES was designed to be used by individual children, hence the fact that ECHOES seems to engender increased social interactions with human partners who are not, strictly speaking, interacting directly with the technology themselves is all the more striking.

The impact that different learning activities may have had on the initiations and responses of children with ASC should be considered. Although the complexity of tasks may have increased over time, some actions such as those involving objects may have necessitated little or no communication with a social partner. As these actions have not been counted as initiations, or responses, the increase in ASC children's initiations and responses cannot be explained by the nature of the activities alone. This conclusion is further supported by that fact that the TD children experienced roughly the same sequence of activities and showed little change over time relating to the increase in the complexity of activities undertaken. A more plausible explanation may be that some activities may be more likely to provoke more initiations or responses to a social partner than others. The impact of particular activities is something that has been proposed for further analysis by [Alcorn 2016] who has also explored the features that may contribute to this. Another possible explanation might be that the children became familiar with the set-up, the technology and the researchers, which led to the decrease in their anxiety and an increase in their motivation to communicate with the social partners. However, given that we have taken special care to familiarise the children with all of the researchers prior to the intervention specifically to reduce children's anxiety (see Fig. 6, Section 4 for the discussion of the procedure), it is less likely that the increase in initiations can be explained by the decrease in anxiety and rise in the familiarity of the set-up.

If this were the case, then we would also expect significant improvements on the post-test, as by that point the children were very familiar with both the activities and the researchers. Such improvements at post-test were not observed.

7.2 Evidence of transfer and impact on schools

Whilst post-test quantitative analysis suggests no evidence of transfer from the ECHOES environment to the table-top activities, it is possible that further analysis of other classroom and play contexts may have provided more supportive evidence. Although the table-top activities defined in this study were chosen carefully on the basis of existing literature and other similar empirical studies, they may not in fact correspond to children's typical activities. However, each practitioner/researcher working with a child using ECHOES was asked to keep a research diary in which observations of the child's behaviour and reactions to the system were recorded. At the sites where the practitioner was a 'support for learning', or classroom assistant based in the school, the research diary also charted any changes or improvements seen in the classroom. Qualitative data collected through these research diaries, as well as through observation, teacher reports and interviews, suggests that there is potential for transfer from ECHOES to the classroom environment.

Classroom assistants who sat in on ECHOES sessions reported that they were able to observe specific behaviours in some children that would not have been apparent without the ECHOES environment. This provided insights to children on an individual level. By having a focussed exploration environment, staff were able to observe children benefitting from success within an educational context, and in several cases transferring this to the classroom. Several examples are given here, for illustrative purposes:

a. One child had consistently had difficulty in initiating the end phase of interactions or classroom work. Within ECHOES the signing gestures of the virtual character included "my turn", "your turn" and "finished", accompanied by the corresponding phrases being spoken by Andy (the "finished" sign was accompanied be the spoken phrase "all done"). Following interactions with ECHOES, on several occasions and without being prompted, the child indicated to the researchers that he wished to finish working with ECHOES. The child continued to indicate completion in the class (i.e. outside of the ECHOES context), using a combination of verbalisation and signing, along with his prompt strip. This extended beyond the classroom, with examples provided by teachers of similar actions in the playground and outside of school. "We were able to identify . . . the signs he was giving and then how to support them." (Class teacher). This new ability also resulted in the reduction in this child's stress levels, as reported by the teachers.
b. A child who initially appeared to be lacking in confidence and did not attempt to interact with the researchers became much more verbal, responding to researchers and even initiating conversations by the end of the study. For example, he initiated communication by waving to researchers across the playground and engaging in conversation with them about a book he was reading, elaborating on points of conversation and providing new information unprompted: this suggested a potential growth in confidence, both in his use of the system and his interactions with the researchers.
c. Another child began the study showing limited verbalisation and little acknowledgement of the researchers' presence. However, as he became more fascinated with the bubbles in the initial ECHOES activity, he began to request to return to this scenario. Initially these were short utterances, such as "bubbles"; eventually these became complete sentences (supplemented by the prompt strip symbols) such as "I want bubbles please". This change was remarkable and with the researchers responding to these requests his challenging behaviours diminished considerably by the end of the study.

d. Two boys, who previously played basketball individually in the playground, started taking turns at throwing their basketball through the hoop. This occurred without any facilitation from school staff and was observed on a number of occasions both during and after the ECHOES study. This particular example of turn taking could be mapped to an activity in the ECHOES system where the child takes turns with Andy to throw balls through a cloud (to make them change colour).

e. A child who showed little awareness of the researchers initially began to verbalise his requests for his favourite scenarios (facilitated by a prompt strip). It was noticed that over-exaggerated praise (clapping and cheering) had a much greater effect than feedback in the ECHOES system. The child began to verbalise much more frequently and his eye contact improved. The transfer of knowledge of the need for exaggerated praise to the classroom resulted in increased confidence in tasks such as writing or answering questions, greater awareness of his achievements, and attempts to repeat them and to share them with others. "This is a boy that now expects us to give him a round of applause when he achieves it and what I loved about it is the interaction he then had with the adults . . . wanted us all to clap." (SLA).

In addition to the task-related initiations observed, there is anecdotal evidence in later sessions with ECHOES of additional initiations from several children who showed no initial interest in Andy, but who were later observed to greet Andy spontaneously through gestures as well as verbally, including waving and saying "Hi Andy!" when he walked onto the screen. As noted earlier, such behaviours were very surprising to teachers and support workers within the schools, who believed many of the children in question to be non-communicative. One child started spontaneously greeting staff in the classroom. A number of teachers reported increased confidence in several children over the course of the research, both when interacting with ECHOES and in the classroom.

Further description and analysis of such qualitative data is outside the remit of this paper. However, one issue which merits further investigation is the extent to which the positive qualitative outcomes can be attributed to the use of ECHOES, to the presence of technology more generally, or perhaps even to a combination of ECHOES and the changed and highly individualised context in which children found themselves. Individual comments from teachers suggest that the AI agent was a key factor for many children, for although they may have not consistently interacted with him, they nevertheless 'bought into' him as a character. Further work, some which has already begun (e.g. [Porayska-Pomsta et al. 2013]; see also Section 8), will focus on developing and manipulating the agent technology and will serve to illuminate any evidence in relation to the role that agents may play in provoking ASC children to engage in social communication and interaction.

In relation to the impact on schools, teachers benefited from first-hand experience of doing research and of seeing some children in a different light, discovering hidden abilities and potential of children in the context of their using ECHOES. This allowed the teachers to appreciate further the need for supporting children through different needs and the potential of technology in providing such support. Observing the individual children behaving more spontaneously and communicatively in the context of ECHOES than in the classroom was reported by teachers as the key aspect of the ECHOES environment.

Teachers' testimonies highlight their perspective on the research impact, in terms of three categories:

(1) Impact on children's communication: "The [ECHOES] technology has a massive impact on children involved [. . .], it enabled them to communicate and succeed without any barriers or feeling constrained by everyday teaching. They were actually free to explore and learn and develop by themselves, which was absolutely fascinating to see [. . .] especially with children with very poor communication skills." (Head Teacher, mainstream primary school)

(2) Impact on school practices and training of teachers: ”[ECHOES] inspired the school to look how that technology and pedagogy of learning can help all children who have communication difficulties and in fact all children.” (Head Teacher, special school)
”I have learned such an awful lot [. . .]. I can take what I have seen to other areas in the school [. . .] that will help me do my job better.” (SLA, special school)

(3) Impact on the school’s and children’s aspirations: ”As a result of that the school has really high aspirations for children, so it helps the school and the children develop their own understanding of what learning is.” (Head Teacher, special school).

An unexpected impact was teachers’ enthusiasm for continuing the involvement in further research, despite disruption and additional demands imposed on them. At least three schools are still involved in collaboration with individual ECHOES partners on both funded and student projects (see Section 8). A plausible explanation for this might lie in the openness of the communication between researchers and teachers, involving teachers in the decision-making in relation to the studies (which children should participate, when and how) as well as giving them an active voice in the design of the environment itself throughout the research. Furthermore, the fact that ECHOES provided teachers with an environment in which they could witness children’s previously unseen skills and abilities may also have led to their desire to continue to use ECHOES as a component in the broader learning context.

7.3 Features of ECHOES: lessons learned

Overall, the use of technology, especially of the AI agent, in the realm of ASC is promising. Within the literature there is a focus on transferring skills from the technology environment to practical settings [Rao et al. 2008], which represents an established challenge in ASC [Pole and Morrison 2003]. Despite such transfer not always taking place in practice [Reynhout and Carter 2009], there are a number of successful cases, e.g. [Scattone et al. 2006], who used structured observations to determine the extent of generalisation after using Social Stories [Gray and Garand 1993] to enhance social communication skills. While the extent of generalisation is difficult to measure in a meaningful way due to the small number of participants involved in those studies, the use of observations as a measure of success is encouraging. We believe that the ECHOES study presented in this paper contributes further to the collective knowledge of the types of technologies and their specific uses in classroom contexts, especially those involving blended AI and human interaction, that may bring real benefits to children with ASC, and of the methods for gathering evidence of the efficacy of such approaches.

ECHOES was designed to provide a child-centred environment that encourages exploration and play. Children with ASC were seen to engage with ECHOES, and to interact with the virtual character and the objects therein. As discussed in Section 2.1, during the formative evaluation of the environment, older children with ASC commented on their likes and dislikes, and suggested improvements. A number of design iterations, with input from children, teachers and others with expertise in designing technology for children with ASC, ensured that the resulting environment was engaging and well matched to the target audience of young children with ASC. The data collected demonstrated that the children did engage and interact with the AI agent by addressing comments and questions to Andy, and initiating social communication with him. For example, when Andy failed to put a coloured ball in the matching box, children would often prompt him directly, e.g. by pointing to indicate the correct box. Multiple examples of evidence were collected showing children socially communicating with the virtual character.

The learning activities, delivered through game play, served to provide an environment in which there were many opportunities for children to engage in spontaneous communication behaviours. The approach adopted through the design of the environment of stimulating and prompting interaction appears to be a productive one, and an alternative to interventions in which the locus of control is with the practitioners/adult rather than with the child.

The evidence-based approach, in this case derived from SCERTS, provided an informed basis for the design of learning activities, and for later analysis of the children's social communication behaviours. As SCERTS is also an approach used in a number of schools in the UK and elsewhere, this means that environments such as ECHOES can be integrated into the wider curriculum of the school, and be directly related to specific learning goals and support for each child, regardless of the wide range of individual needs.

The design, formative, and efficacy evaluation stages of the ECHOES project were conducted in schools, rather than laboratories. This required an approach that was carefully planned and orchestrated in consultation and collaboration with schools and teachers, including additional time for classroom observation and familiarisation activities between children and researchers, targeted training of both researchers and practitioners, and flexibility in the final evaluation studies to accommodate the needs of individual children and the specific schools' contexts. Although this meant that the exact conditions of the evaluation study differed across sites (e.g. smaller rooms meant that social partner position in relation to the child varied), children engaged with ECHOES and communicated socially with both human and virtual partners at all sites.

In relation to the approach taken in ECHOES, a major requirement was that the environment should be suitable for use by children with ASC who may also have learning difficulties and limited language abilities. Children with ASC and without intellectual difficulties have more frequently been the target of interventions [Fletcher-Watson et al. 2016], but, arguably, may not have as great a need as those described as so-called "low functioning". Although in the analysis presented in this paper all ASC children are considered together, a future analysis that distinguishes between children based on their intellectual development may reveal greater gains for the latter group.

Children both with and without learning difficulties successfully engaged with ECHOES. In addition to those children whose data was reported in this paper, a number of other children participated in our studies . This included a number of children with very little language and who typically found it difficult to participate in other classroom activities. For all children, sessions with ECHOES were included in their regular (and familiar) classroom schedule planners. On the whole, they participated eagerly in these sessions. In a number of cases they were accompanied by teaching assistants who worked specifically with them. In many cases the teaching assistants commented on their surprise at how much the children were able to do within ECHOES. This was reinforced by observations of classroom teachers, when shown videos of the children demonstrating spontaneous communication behaviours when interacting with ECHOES. The abilities of children with ASC described as "low functioning", and their potential for engaging in communication with others, should not be underestimated.

A major feature of ECHOES, arising from earlier studies [Alcorn et al. 2011], was that it was designed to include the practitioner/researcher as a social partner for the child, in addition to the AI agent. This provided many opportunities for social communication that were frequently taken up throughout the sessions. This demonstrates the potential for technology such as ECHOES to serve as the object of social communication, and as a stimulus or catalyst for spontaneous communication behaviours, as well as (rather than exclusively) an environment within which social interaction can take place. The role of the human partner should be considered further when developing AI technology, and arguably other forms of technology, for use by children with ASC in relation to social communication. The role of the human partner should be explicitly considered, both in the design and evaluation phases, and also in the intended context of use.

7.4 Reflections on the ECHOES' evaluation methodology: Efficacy vs. Effectiveness

In evaluating ECHOES we framed our findings in terms of two important and related societal questions concerning ASC: (i) optimal outcome [Suh et al. 2016], often considered to equate to recovery from autism, and (ii) neuro-diversity. Traditionally, optimal outcomes are sought in intervention studies by looking for pre-/post-test differences [Salkind 2010]. Here, we offer an alternative way of framing the question of optimal outcome, not in terms of recovering from autism, but in terms of changing the environment and finding associated changes in performance and behaviour. This is arguably something that any technology-based intervention is suited to, given the specific and precise nature of its development and potential for fine-grained data on how specific changes in the environment might lead to concomitant changes in users' behaviours and performance. We argue that this reframing has two benefits. First, it forces us to look at the process of any change, rather than looking for high level differences (which can be hard to measure and even harder to show changes in). Second, it means that we have the potential to include more children in our evaluation rather than creating inclusion/exclusion criteria (based on group matching), which might result in fewer children participating in studies.

Rao et al. 2008 argued that before interventions are rolled out into large-scale randomised control trials (in essence, trials that are focused on ascertaining effectiveness), smaller scale efficacy studies should be done to demonstrate their potential. We argue that our evaluation resides at a 'pre-efficacy' stage, in which we believe that ECHOES shows promise in that children with both autism and intellectual difficulties engage with the environment, interact with the AI agent and with a human partner, progress through different learning activities and seem to change their behaviour whilst in the environment. However, our results are tempered with caution, given the lack of a direct comparison group (e.g. an ASC wait-list control). Our findings suggest, though, that the type of technology-based intervention adopted in the ECHOES project may result in real behaviour change, even if it is limited to the ECHOES environment. More recent research based on ECHOES by Alcorn 2016 provides supporting evidence for these conclusions.

Generalisation and ascertaining whether generalisation is maintained on follow-up has been considered to be the 'holy grail' of intervention studies in ASC [Rajendran 2013]. However, our results suggest that another way to reframe the question may be to ask what it is about a particular environment (in this case, the ECHOES environment) that brings about change within that context. If neuro-diversity is the acceptance of people who are neuro-atypical, then these environmental clues may be immensely valuable if we seek to fit the environment to the person, rather than the other way round [Rajendran 2013].

However, such an approach raises questions that are also key to the ongoing debate within evidence-based practice in education as to what constitutes (or should constitute) 'good' evidence in educational research (see e.g. [Biesta 2007, 2013]). This debate highlights a tension between the need to generate evidence that can be trusted, and is on par with the evidence emerging from the biological sciences (and indeed using the same methods, e.g. see [Guldberg 2017; US National Research Council 2002]), and the realities of live classroom environments, where the focus is on individual children's needs, on fine grained changes in their behaviour, and on the front-line practitioners' ability to spontaneously adapt in pedagogically beneficial ways to often unexpected classroom/learning events, [Guldberg et al. 2017; Reichow et al. 2008; Stahmer et al. 2011]. The idiosyncratic nature of autism adds to the challenge of reconciling the need to understand the individual child in the context of specific environments and situations, and the desire to generalise the intervention approaches used along with any evidence of their success. On the whole, 'hard' evidence, i.e. that which is typically generated through randomised control trials, seldom infiltrates real classroom practice, not least because interventions that are tightly controlled in the clinical or lab environments are notoriously hard to replicate by clinically untrained front-line practitioners in often messy and only partly predictable classroom settings, e.g. [Dingfelder and Mandell 2011; Parsons et al. 2013; Porayska-Pomsta 2016].

8 CONCLUSIONS AND FURTHER DIRECTIONS FOR RESEARCH

This paper presented an evaluation of the ECHOES technology-based intervention, which aimed to ascertain its potential to facilitate autistic children's ability to engage in social interaction. The evaluation was conducted in several schools across the UK, and focused on working with the so-called "low functioning" autistic children across the 4-14 age range. The results of the evaluation showed a significant increase in the proportion of children's responses to the human social partners, and suggested positive trends with respect to children's initiations to both social partners (an AI agent and a human practitioner). These results are very encouraging in their own right. However, we also believe that the results highlight a number of important considerations for further research directions in the area of technology for autism education, and on the role that technology, especially AI-based technology, may play in helping us understand and respond to children with ASC, and in informing technology-enhanced educational practices more broadly.

A key aspect of the approach presented in this paper was the way in which the intervention was implemented in the different school contexts, placing an emphasis on ecological validity in assessing the educational efficacy of the ECHOES approach, as well as on the need for flexibility in the way that the AI technology's role is prioritised and understood by the practitioners and researchers. Specifically, an over-reliance on technological support alone may be misguided in autism related practices and, arguably, in broader educational contexts as well, where it can lead to many important opportunities for learning being either ignored or missed altogether. In ECHOES, allowing a human practitioner to provide social partnership on demand alongside the AI agent has led to: (i) the revelation and enhancement of many individual children's specific abilities, (ii) consideration of how research and educational practice may be consolidated, and (iii) rethinking of how technology may be designed to perform an optimal role in supporting both the learners and educational practitioners in achieving the desired learning outcomes. Teachers who participated in the research were able to tailor their classroom support to the individual children based on their newfound understanding of what the children could do, rather than on what the children had revealed about themselves to date in typical classroom situations.

However, there are limitations associated with the study, many of which are artefacts of working in the context of autism – a condition characterised by high individual heterogeneity and co-morbidity with other developmental conditions. These limitations include the fact that, although a relatively large number of participants took part in this research as compared to other studies involving autistic children, the number coupled with the idiosyncratic differences between the individual participants means that the results may not generalise to all contexts and all individuals beyond this research. Furthermore, as indicated throughout the paper, the lack of a comparable control group and activities means that the conclusions drawn from this research apply mainly to within-the-environment improvements. These limitations are being addressed in further work, some of which has already began or has been undertaken.

Since the completion of the ECHOES project, many of the insights gained have served to inspire new ways of thinking by both the individual ECHOES researchers and by the teachers who were involved in ECHOES, with several schools having become long-term partners in ongoing research. For example, the SHAPE project [Guldberg et al. 2017; Parsons et al. 2015] focused explicitly on the challenge of bridging between research goals and outcomes, and school needs and practices. SHAPE investigated how different technologies developed for children with ASC may be embedded in different schools' everyday practices to reveal any mismatches and commonalities in teachers' vs. researchers' perspectives vis-á-vis what children might actually find inspiring, motivating and useful, and to highlight any opportunities for pedagogical and technological innovation. This focus was further extended in another project, SHARE-IT [Porayska-Pomsta et al. 2013], which aimed to sustainably include the perspectives of all stakeholders concerned (parents, teachers, researchers and children), for example through allowing the teachers and the parents to (re-)configure both the children's profiles and the specific activities within the environment, thus supporting the construction and periodic modification of the (user) child models underpinning the system's interactions with the children. SHARE-IT's focus reflected the insights gained via the ECHOES project, specifically – the need to increase the range and variability of the learning activities and provide more customisable and malleable technologies in terms of their designs and deployment in different contexts. Crucially, through its further investment in the planning technology that underpinned the ECHOES agent's behaviours, SHARE-IT also aimed to respond explicitly to the debate that was on-going within the ECHOES project with respect to the degree and the nature of artificial intelligence that was actually needed to provide the necessary support. Although SHARE-IT resulted in a much more socially able AI agent than was achieved during ECHOES, this debate is still on-going and the questions it raises remain open.

Post ECHOES and the projects that follow from it, it is clear that there is a need for balance between technological innovation push and the educational needs pull to allow us to consider how technology may be employed optimally in the context of autism-focused education. There is a definitive need for flexibility in the way that the respective roles of technology and human social partners are understood and managed, with the best-case scenario, i.e. one that serves the learning process, being that technology and the human practitioner enhance, rather than override, each other. The question of flexibility of the environment within which children are motivated to spontaneously engage in social interaction with others and the specific opportunities that technology provides with respect to supporting children in doing so was the subject of a PhD thesis by Alcorn 2016 which was inspired by the ECHOES project and which utilised its data as a starting point. Alcorn's investigation focused on the motivational potential of subjectively inconsistent (i.e. discrepant, unintentional and non-designed) aspects in game-like virtual contexts for young children with ASC. Her analysis of the ECHOES video data illustrated that a heterogeneous group of children all reacted frequently and socially to naturally occurring discrepant aspects within ECHOES. This led to the creation of a set of high-level design principles that might facilitate similar patterns of spontaneous, positive initiations around discrepancies. These design principles were implemented in a set of new touch-screen games that sought to establish, and then deliberately violate, child expectations. The results of Alcorn's research suggest that it is possible to motivate children's communication – specifically their initiations – by including deliberately designed discrepancies in the technology.

While the interdisciplinary nature of the ECHOES project presented the team with many intellectual challenges, it also offered a richness of insight that is now bearing fruit in many follow-on projects as just illustrated. The ECHOES project and the outcome of the use of ECHOES technology as reported in this paper contributed to a change in the team's perception, not only of what is technologically possible, but also of what needs to be approached with an open mind. In particular, if the support offered is to go beyond paying lip service to the question of neuro-diversity and inclusion, there is a need for balance between technology, in this case AI technology, and the human role in delivering support to the child, and in moulding the learning environment to the needs of individual children. We believe that the results presented in this paper, along with the way that the ECHOES technology was designed and deployed, provide an evidence-based starting point for challenging and addressing the existing orthodoxies in relation to the perceived and real potential of AI technology for autism education and research.

1. The activity opens with the non-interactive garden background (trees and hills), the interactive cloud and a numerical counter (displaying zero). Ambient garden sounds, such as chirping birds, play in the background.
2. Andy enters, draws the child's attention to the cloud by saying "Look, a cloud, let's make it rain!" and jumps up to shake it.
3. A rain animation and sound effect are triggered, causing a small flower to grow from the ground, and incrementing the counter.
4. Andy draws the child's attention to the cloud-flower relationship by turning and exaggeratedly looking from the cloud to the flower. After a pause, he re-orients to the front and looks out 'at' the child, and invites him to take a turn.
5. The child can 'shake' the cloud by touching it and rapidly moving it from side to side. The child can move the cloud across the screen to any desired location using touch-and-drag.
6. The flowers can be picked (using an upward dragging movement) and moved around the screen; wiggling the flower heads makes them sway.
7. When the child grows a flower, Andy provides positive feedback.
8. Continuing to shake the cloud increases the size of a flower. Shaking and moving the cloud across the screen grows a whole line of flowers.
9. When the child stops shaking the cloud, and Andy has given feedback, he then announces "My turn" and proceeds to act.

Fig. 10. Example of Flower Growing activity.

9 REFERENCES

A SUPPLEMENTARY MATERIALS

A.1 Examples of ECHOES Activities and child-ECHOES interaction

In order to give readers a sense of the interaction engendered by ECHOES, we describe two contrasting learning activities. These activities highlight the types of actions that are available to the child, to Andy, and to the researcher/practitioner (via the GUI), and show how these actions relate to the socio-communicative behaviours targeted in ECHOES.

The first example, flower growing (Fig. 10), involves shaking the interactive cloud to make it rain in order to grow flowers, and is one of the simplest activities in ECHOES. The activity is exploratory in nature; it is about fostering playful exploration and immersion in the environment. Many children were captivated by the flower growing and would shake the cloud for long periods without a break, appearing to enter into a flow-like state. Such interactions can continue indefinitely, as the activity has no explicit end point. Although it could be argued that such a set up might have encouraged obsessive behaviours in children, in ECHOES it was used explicitly to foster children's sense of calm. In line with the ethos of SCERTS, positive emotional states, especially a feeling of calm, provide a basis that is conducive to social interaction. During the periods of flow-like states, the rain sound effects would continue. Here Andy's role was to periodically give positive feedback such as "Wow!" or "Cool!", whereas the role of the human practitioner was to manage the transitions between activities, by carefully gauging the child's state. The on-screen flower counter was a useful tool for the practitioner in helping them to create a goal or a predictable end-point, preparatory to transitioning

1. Three coloured boxes (red, blue, and yellow) and nine bouncy balls, three of each colour appear at the start of the activity.
2. Andy enters and identifies the task, saying ``Let's tidy up these balls into the boxes''.
3. Andy models the ball-sorting action by walking to his target ball, and then dropping it into the box of the same colour.
4. Andy then turns and looks 'to' the child and invites him or her to take a turn. He indicates a specific ball through one of: gaze only, distal pointing (with arm and forefinger extended), or contact pointing, walking to the ball and touching it with his forefinger.
5. Children sort balls by touching one and dragging it up and around the boxes, dropping it in the top. Balls cannot be dragged across boxes, and can only go in the box of the same colour.
6. If incorrectly sorted, balls roll off the top and bounce to the floor.
7. When a box is filled, the child receives an attractive sensory reward. The red box and balls vanish to be replaced by three bubbles. The blue box becomes a fireworks display, and the yellow box yields three cartoon bumblebees, which buzz around the screen for several seconds before vanishing.

Fig. 11. Example of Sorting Balls into Boxes activity.

to a new activity. It also provided very clear opportunities for the practitioners to initiate bids for interactions in ways that were contextually relevant to the children. For example, the practitioner might say ”Look, you grew 12 flowers! You can grow three more and then it is time for a new game”. This type of activity offers an important example for (a) how the AI and human intelligence and skills were blended in ECHOES and (b) why it is important to allow for such blending to take place in order to build on the specific strengths of both AI and human understanding of the child and the possible support.

The second example, sorting balls into boxes (Fig. 11), is a very different type of activity to the flower growing one, as it is goal-oriented, involving a sequence of steps toward a clearly identifiable end goal. Teachers suggested this activity during one of the design sessions, having observed that many children with ASC enjoy sorting objects or helping to tidy the classroom. Sorting objects is also a typical activity used in autism intervention as it provides ample opportunities for modelling and practicing turn taking and joint attentional skills.

Having multiple objects with which the child can interact allows Andy to direct the child's attention to those objects explicitly, providing an opportunity for the child to practice following referential pointing gestures and gaze. As before, the practitioner can use the GUI to make Andy repeat his prompting as needed. Children can also act by giving Andy a ball to sort by dragging it to him and holding it anywhere over his body, until he accepts and says ”Thanks. I'll put it in the box”. Some children discovered this action independently, but in other cases the practitioner had to draw the child's attention to this possibility, sometimes modelling the "giving” action for the child.

Even though ball sorting is much more structured than flower growing, it is important to emphasise its inherent flexibility. Balls can be sorted and boxes filled in any order, even if some children independently identified and followed a particular order (e.g. completing boxes one-at-a-time, left to right). The boxes vanish upon activity completion, providing an explicit reward, such as a release of fireworks, bumble bees or poppable soap bubbles. This gives clear cues to the child that they have completed the activity and it signals to them that they have been successful.

In all activities, Andy reliably gives positive verbal feedback for the child's actions. To this end he may use any of several phrases, such as "Good job", "Cool" and "Woah", all of which are supported by appropriate gestures such as Andy extending his arms in the air or using a thumb up, the latter also being used in the Makaton language to mean "Well done". Andy's planning architecture prioritises positive feedback over other possible actions. The overall pacing of these actions is slow compared to what adults may initially expect, or what may be customarily designed in programmes for TD children, but this pacing gives children with ASC time to notice and process what is happening, particularly where it involves Andy directing attention or giving instructions.

Many children were able to take turns with Andy, either spontaneously or when given additional support by the practitioner (e.g. explicitly drawing the child's attention to the potential of turn-taking or asking him to consider whose turn it should be next). However, even where children agreed that they should take turns or it was Andy's turn next, they often found self-inhibition difficult, likely due to developmental age. Almost all children needed repeated prompting to wait while Andy took a turn. Aside from its importance to turn taking as a social skill, waiting was especially important in ECHOES due to the functioning of the agent's planner. For example, if the child began manipulating the cloud when it was supposed to be Andy's turn, Andy would then re-plan and give feedback, rather than take his turn, which sometimes resulted in higher than expected latency in his reactions. This latency was inherent in the planning architecture chosen and as such constitutes an interesting challenge for AI technologies in real-time applications such as ECHOES more generally. Nevertheless, this inconsistency between what the children were told to expect and what occurred in the environment was potentially confusing for some of them. In situations where delays occurred, the practitioner could use dedicated buttons on the GUI to override the planner and to make Andy take a turn on demand. Other direct instructions included the command for Andy to repeat instructions/prompting, or to leave when it was time to end the activity.

When transitioning to a new activity, the practitioner would trigger the transition process from the GUI: in such cases Andy would walk out, and a red bubble would appear in the centre of the screen, slowly floating off to the right to signal the end of the current activity. This way of signalling transition between activities was the same in all of the ECHOES activities, and gave a clear visual cue that the activity was ending. To indicate the beginning of a new activity, a green 'transition' bubble would float on the screen from the left, shortly followed by Andy entering the scene. Again this signalling was used across all activities.

An important point to take away from both of these activity descriptions, and one that directly impacts the reporting of the results, is that the number of initiations (to which a child could potentially respond) varies across activities, sessions, and individuals due to the agent's reactive planning and the necessity for occasional interventions by the human researcher/practitioner (e.g. making the character repeat instructions). Thus, child responses to partner initiations must be compared as proportions, rather than as raw numbers, as explained in detail in Section 4.

A.2 ECHOES system architecture

The ECHOES environment comprises three distinct software components, which communicate with one another to detect and process user actions and to select appropriate responses. These components include:

(1) Multimodal Fusion Engine (MFE), which combines low-level input events (e.g. touch-screen input) into higher-level composite multimodal events. The composite events allowed in ECHOES include gaze and eye tracking

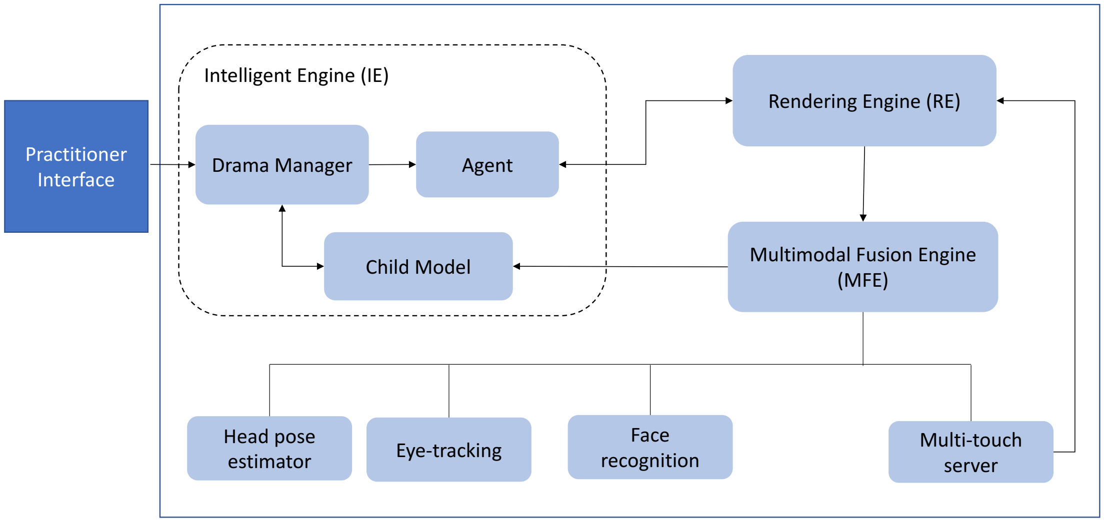

Fig. 12. ECHOES overall system architecture.

data, multi-touch events as well as face recognition. The MFE utilises the ICE middleware to allow for an open-source package with an active development and support community; industry-standard application development; use in a wide set of operating systems and programming languages; publish-subscribe messaging and direct module-to-module communication; and use of structured, strongly typed messages.

(2) Intelligent Engine (IE), which selects actions for the virtual character and specifies changes to the state of the world based on the current learning objectives and the child's behaviours as represented by the composite events that are sent by the MFE. The composite events detected by the environment are recorded in the Child Model which is used as the basis for the determining the sequencing and duration of the learning activities within ECHOES. The sequencing and the duration of the activities within ECHOES can also be determined by the practitioner, through the Practitioner Interface, which is connected directly to the Drama Manager. The practitioner interface proved necessary to further ensure that the intervention/pedagogical decisions made during children?s interactions with ECHOES were beneficial to the children, and to account for the fact that in the real-school (i.e. non-lab) contexts in which the studies reported have been conducted only the multi-touch events provided a consistently robust source of information about children's actions. Whilst this was a hindrance from the point of view of automating all of the decision-making processes within ECHOES, it is not an unusual difficulty encountered when advanced sensing technologies are used in the wild. With respect to the target population discussed in this paper coupled with the use of the technology in physically non-restricted ways (e.g. children were not made to sit at a particular distance from the screen), the sensing technology did not prove up to the task, even though it worked robustly during lab tests. The key component of the IE is the agent which is underpinned with an emotional planning architecture called FAtiMA. This architecture allows the agent plan and

execute its actions based on the information from other components in the system in a way that is both reactive in real-time and deliberative over the duration of each ECHOES activity. The emotional element is critical as it allows the agent's decisions to be driven by its emotional thresholds. Whilst in ECHOES these thresholds were set to low for positive emotions and high for negative emotions (Andy is a positive social partner), the architecture offers the flexibility for nuancing the emotional displays of the agent as needed by the different contexts, activities and user cohorts.

(3) Rendering Engine, which modifies its display and behaviour as necessary, based on the actions requested by the IE for the virtual character and on the world updates sent by the MFE.

Further details of the ECHOES' software architecture and individual components are given in [Foster et al. 2010] and [Bernardini and Porayska-Pomsta 2013].

ACKNOWLEDGMENTS

This research was funded by the ESRC and EPSRC under TLRP-TEL programme (grant number RES-139-25-0395-A). We would like to thank all the participating schools, including children, teachers and parents, Michael Anderson for providing the voice of the AI agent, as well as TLRP-TEL programme director Prof Richard Noss for his invaluable advice and encouragement throughout the project.

REFERENCES

B. Abirached, Y. Zhang, J. Aggarwal, B. Tamersoy, T. Fernandes, J. Miranda, and V. Orvalho. 2009. Improving communication skills of children with ASDs through interaction with virtual characters.. In *IEEE 1st International Conference on Serious Games and Applications for Health (SeGAH)*. 1–4.

Alyssa Alcorn, Helen Pain, Gnanathusharan Rajendran, Tim Smith, Oliver Lemon, Kaska Porayska-Pomsta, Mary Ellen Foster, Katerina Avramides, Christopher Frauenberger, and Sara Bernardini. 2011. Social communication between virtual characters and children with autism. In *Proceedings of the 15th international conference on Artificial intelligence in education (AIED'11)*. Springer-Verlag, Berlin, Heidelberg, 7–14.

A. M. Alcorn. 2016. *Embedding novel and surprising elements in touch-screen games for children with autism: creating experiences 'worth communicating about'*. Ph.D. Dissertation. University of Edinburgh, Edinburgh, Scotland.

K. Anderson, E. André, T. Baur, S. Bernardini, M. Chollet, E. Chryssafidou, I. Damian, C. Ennis, A. Egges, P. Gebhard, H. Jones, M. Ochs, C. Pelachaud, K. Porayska-Pomsta, P. Rizzo, and N. Sabouret. 2013. The TARDIS framework: Intelligent virtual agents for social coaching in job interviews. In *Advances in Computer Entertainment*. Springer, 476–491.

A. Anwar, M. Rahman, S. Ferdous, S. Anik, and S. Ahmed. 2011. AA computer game based approach for increasing fluency in the speech of the autistic children.. In *Proceedings of the 11th IEEE International Conference on Advanced Learning Technologies (ICALT '11)*. 17–18.

APA. 2013. *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, Text Revision (DSM-5)*. American Psychiatric Association.

R. Aylett, N. Vannini, E. Andre, A. Paiva, S. Enz, and L. Hall. 2009. But it was another country: agents and intercultural empathy.. In *8th International Conference on Autonomous Agents and Multiagent Systems (I/International Foundation for Autonomous Agents and Multiagent Systems.)*, Vol. 1. 329–336.

R S. Baker. 2016. Stupid tutoring systems, intelligent humans. *International Journal of Artificial Intelligence in Education* 26, 2 (2016), 600–614.

E. Barakova, G. van Wanrooij, R. van Limpt, and M. Menting. 2007. Using an emergent system concept in designing interactive games for autistic children.. In *Proceedings of the 6th international conference on Interaction Design and Children. (IDC '07)*. ACM, 73–76.

L. Barnard-Brak, A. Brewer, S. Chesnut, D. Richman, and A. M. Schaeffer. 2016. The Sensitivity and Specificity of the Social Communication Questionnaire for Autism Spectrum with Respect to Age. *Autism Research* 9, 8 (2016), 838–845.

L. Bartoli, C. Corradi, F. Garzotto, and M. Valoriani. 2013. Exploring Motion-based Touchless Games for Autistic Children?s Learning.. In *Proceedings of the ACM Interaction Design and Children Conference (IDC ' 13)*. ACM, New York, USA.

A. Battocchi, F. Pianesi, D. Tomasini, M. Zancanaro, G. Esposito, P. Venuti, A. Ben Sasson, E. Gal, and P. L. Weiss. 2013. Collaborative puzzle game: a tabletop interactive game for fostering collaboration in children with autism spectrum disorders (ASD).. In *Proceedings of the ACM International Conference on Interactive Tabletops and Surfaces. (ITS '09)*. ACM, 197–204.

N. Bauminger. 2002. The Facilitation of Social-Emotional Understanding and Social Interaction in High-Functioning Children with Autism: Intervention Outcomes. *Journal of Autism and Developmental Disorders* 32 (2002), 283–298.

R. Beaumont and K. Sofronoff. 2008. A multi-component social skills intervention for children with asperger syndrome: The junior detective training program. *Journal of Child Psychology and Psychiatry* 49 (2008), 743–753.

S. Bernardini and K. Porayska-Pomsta. 2013. Planning-based social partners for children with autism.. In *Proceedings of the Twenty Third International Conference on Automated Planning and Scheduling (ICAPS-13)*. AAAI Press, Rome, Itay.

S. Bernardini, K. Porayska-Pomsta, and Smith T. J. 2014. Echoes: An intelligent serious game for fostering social communication in children with autism. *Information Sciences* 264 (2014), 41–60.

G. Biesta. 2007. Why ?what works? won?t work: Evidence-based practice and the democratic deficit in educational research. *Educational Theory* 57, 1 (2007), 1–22.

G. Biesta. 2013. *The Beautiful Risk of Education.* Boulder and London: Paradigm Publishers.

A. Bosseler and D. Massaro. 2003. Development and evaluation of a computer- animated tutor for vocabulary and language learning in children with autism. *IJournal of Autism and Developmental Disorders* 33, 6 (2003), 653–672.

M. O. Cavazza, F. Charles, and S.J. Mead. 2002. Character-based interactive storytelling. *IEEE Intelligent Systems* 17, 4 (2002), 17–24.

H.-M. Chiang and M.. Carter. 2008. Spontaneity of communication in individuals with autism. *Journal of Autism and Developmental Disorders* 38, 4 (2008), 693–705.

S. Cobb. 2007. Virtual environments supporting learning and communication in special needs education. *Topics in Language Disorders* 27, 3 (2007), 211–225.

K. Dautenhahn and I. Werry. 2004. Towards interactive robots in autism therapy: Background, motivation and challenges. *Pragmatics and Cognition* 12, 1 (2004), 1–35.

M. Dennis, D. J. Francis, P. T. Cirino, R. Schachar, M. A. Barnes, and J. M. Fletcher. 2009. Why IQ is not a covariate in cognitive studies of neurodevelopmental disorders. *Journal of the International Neuropsychological Society* 15, 3 (2009), 331–343.

J. Dewey. 1998. *Experience and Nature.* Dover Publications Inc.

João Dias and Ana Paiva. 2005. Feeling and Reasoning: A Computational Model for Emotional Characters. In *Progress in Artificial Intelligence*. Lecture Notes in Computer Science, Vol. 3808. Springer Berlin, Heidelberg, 127–140.

H E. Dingfelder and D S. Mandell. 2011. Bridging the research-to-practice gap: an application of diffusion of innovation theory. *Journal of Autism and Developmental Disorders* 41, 5 (2011), 597–609. https://doi.org/10.1007/s10803-010-1081-0

W. Farr, N. Yuill, and H. Raffle. 2010. Social benefits of a tangible user interface for children with autistic spectrum conditions. *Autism* 14, 3 (2010), 237–252.

S. L. Finkelstein, A. Nickel, L. Harrison, E. A. Suma, and T. Barnes. 2009. cMotion: A new game design to teach emotion recognition and programming logic to children using virtual humans.. In *Proceedings of the 2009 IEEE Virtual Reality Conference.* 249–250.

S. Fletcher-Watson. 2014. A targeted review of computer-assisted learning for people with autism spectrum disorder: Towards a consistent methodology. *Review Journal of Autism and Developmental Disorders* 1, 2 (2014), 87–100.

S. Fletcher-Watson, H. Pain, S. Hammond, A. Humphry, and H. McConachie. 2016. Designing for young children with autism spectrum disorder: A case study of an iPad app. *International Journal of Child-Computer Interaction* 7 (2016), 1–14.

M. E. Foster, K. Avramides, S. Bernardini, J. Chen, C. Frauenberger, O. Lemon, and K. Porayska-Pomsta. 2010. Supporting children?s social communication skills through interactive narratives with virtual characters.. In *Proceedings of ACM Multimedia*. ACM, Florence, Italy.

C. Frauenberger, J. Good, A. Alcorn, and H. Pain. 2013. Conversing through and about technologies: design critique as an opportunity to engage children with autism and broaden research(er) perspectives. *International Journal of Child-Computer Interaction* 1, 2 (2013), 38–49.

C. Frauenberger, J. Good, and W. Keay-Bright. 2010. Phenomenology, a framework for participatory design. In *Proceedings of the 11th Biennial Participatory Design Conference (PDC '10).* ACM, New York, NY, USA, 187–190.

C. Frauenberger, J. Good, and W. Keay-Bright. 2011. Designing technology for children with special needs: bridging perspectives through participatory design. *CoDesign* 7, 1 (2011), 1–28.

O. Golan, E. Ashwin, Y. Granader, S. McClintock, K. Day, V. Leggett, and S. Baron-Cohen. 2010. A multi-component social skills intervention for children with asperger syndrome: The junior detective training program. *Journal of Autism and Developmental Disorders* 40 (2010), 269–279.

M. Goodwin. 2008. Enhancing and accelerating the pace of autism research and treatment: The promise of developing innovative technology. *Focus on Autism and Other Developmental Disabilities* 23, 2 (2008).

C. Gray and J. Garand. 1993. Social Stories: Improving Responses of Students with Autism with Accurate Social Information. *Focus on Autism and Other Developmental Disabilities* 8, 1 (1993), 1–10.

Ouriel Grynszpan, Jean-Claude Martin, and Jacqueline Nadel. 2008. Multimedia interfaces for users with high functioning autism: An empirical investigation. *International Journal on Human-Computer Studies* 66, 8 (2008), 628–639.

O. Grynszpan, P. L. T. Weiss, F. Perez-Diaz, and E. Gal. 2014. Innovative technology-based interventions for autism spectrum disorders: A meta-analysis. *Autism* 18, 4 (2014), 851–871.

K. Guldberg. 2017. Evidence Based Practice in autism educational research: can we bridge the research and practice gap? *Oxford Review of Education* 43, 4 (2017).

K. Guldberg, S. Parsons, K. Porayska-Pomsta, and W. Keay-Bright. 2017. Challenging the knowledge transfer orthodoxy: knowledge co-construction in technology enhanced learning for children with autism. *British Educational Research Journal* 43, 2 (2017), 384–413.

M. Hauck, D. Fein, L. Waterhouse, and C. Feinstein. 1995. Social initiations by autistic children to adults and other children. *Journal of Autism and Developmental Disorders* 25 (1995), 579–595.

G. Herrera, F. Alcantud, R. Jordan, A. Blanquer, G. Labajo, and C. De Pablo. 2008. Development of symbolic play through the use of virtual reality tools for children with autism spectrum disorders. Two case studies. *Sage Publications and The National Autistic Society* 12, 2 (2008), 143–157. https://doi.org/10.1177/1362361307086657

G. Herrera, Casas X., J. Sevilla, L. Rosa, C. Pardo, J. Plaza, R. Jordan, and S. Le Groux. 2012. Pictogram Room: Natural Interaction Technologies to Aid in the Development of Children with Autism. *Annuary of Clinical and Health Psychology* 8 (2012), 39–44.

S. Holt and N. Yuill. 2017. Tablets for two: how dual tablets can facilitate other-awareness and communication in learning disabled children with autism. *International Journal of Child-Computer Interaction* 11 (2017), 72–82.

I. M. Hopkins, M. W. Gower, T. A. Perez, D. S. Smith, F. R. Amthor, and et al. 2011. Avatar assistant: Improving social skills in students with an ASD through a computer-based intervention. *Journal of Autism and Developmental Disorders* 41, 11 (2011), 1543–1555.

J.P. Hourcade, S.R. Williams, E.A. Miller, K.E. Huebner, and L.J. Liang. 2013. Evaluation of tablet apps to encourage social interaction in children with autism spectrum disorders.. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI ' 13)*. ACM, New York, USA, 3197–3206.

N. Kandalaft, M R.and Didehbani, D C. Krawczyk, T T. Allen, and S B. Chapman. 2013. Virtual Reality Social Cognition Training for Young Adults with High-Functioning Autism, Journal of Autism and Developmental Disorders. *Journal of Autism and Developmental Disorders* 43, 1 (2013), 34–44.

P. Kenny, H. Arno, J. Gratch, W. Swartout, D R. Traum, S. Marsella, and D. Piepol. 2007. Building Interactive Virtual Humans for Training Environments. In *Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)*.

H. Kozima, M. Michalowski, and C. Nakagawa. 2009. Keepon: A playful robot for research, therapy, and entertainment. *International Journal of Social Robotics* 1, 1 (2009), 3–18.

C. Lord, M. Rutter, P. DiLavore, S. Risi, K. Gotham, and S. Bishop. 2012. *Autism diagnostic observation schedule: ADOS-2.* Western Psychological Services Torrance.

L. Malinverni, J. Mora, V. Padillo, A. Hervas, and N. Pares. 2014. Pico?s Adventure: A Kinect Game to Promote Social Initiation in Children with Autism Spectrum Disorder. In *Proceedings of the ITASD international conference (IDC ' 16)*. ACM, Paris, France.

D. Massaro. 2006. Embodied agents in language learning for children with language challenges.. In *Proceedings of the 10th international conference on Computers Helping People with Special Needs. (ICCHP?06)*. Springer-Verlag, Berlin, Heidelberg, 809–816.

J. Mora-Guiard, C. Crowell, N. Pares, and Heaton. P. 2016. Lands of Fog: Helping Children with Autism in Social Interaction through Full-Body Interactive Experience.. In *Proceedings of the ACM Interaction Design and Children Conference (IDC '16)*. ACM, Manchester, UK.

L. Mottron. 2004. Matching strategies in cognitive research with individuals with high-functioning autism: Current practices, instrument biases, and recommendations. *Journal of autism and developmental disorders* 34, 1 (2004), 19–27.

P. Mundy, C. Delgado, J. Block, M. Venezia, A. Hogan, and J. Seibert. 2003. *Early social communication scales (escs).* Coral Gables, FL: University of Miami.

NAS 2016. What is autism? Retrieved November 12, 2016 from http://www.autism.org.uk/about/what-is.aspx

A. Ortony, G L. Clore, and A. Collins. 1988. *The Cognitive Structure of Emotions.* Cambridge University Press.

S. Parsons, T. Charman, R. Faulkner, J. Ragan, S. Wallace, and K. Wittemeyer. 2013. Commentary ? bridging the research and practice gap in autism: The importance of creating research partnerships with schools. *Autism* 17, 3 (2013), 268–280.

S. Parsons and S. Cobb. 2011. State-of-the-art of Virtual Reality technologies for children on the autism spectrum. *European Journal of Special Needs Education* 26, 3 (2011), 355–366.

S. Parsons, K. Guldberg, A. MacLeod, G. Jones, A. Prunty, and T. Balfe. 2007. International review of the evidence on best practice in educational provision for children on the autism spectrum. *European Journal of Special Needs Education* 26, 1 (2007), 47–63.

S. Parsons, K. Guldberg, K. Porayska-Pomsta, and R. Lee. 2015. Digital Stories as a method for evidence-based practice and knowledge co-creation in technology-enhanced learning for children with autism. *International Journal of Research and Method in Education* 38, 3 (2015), 247–271. https://doi.org/10.1080/1743727X.2015.1019852

S. Parsons and P. Mitchell. 2017. The potential of virtual reality in social skills training for people with autistic spectrum disorders. *Journal of Intellectual Disability Research* 46, 5 (2017), 430–443.

R. C. Pennington. 2010. Computer-assisted instruction for teaching academic skills to students with autism spectrum disorders: A review of literature. *Focus on Autism and Other Developmental Disabilities* 25, 1 (2010), 239–248.

C. Pole and M. Morrison. 2003. *Ethnography for Education.* Open University Press.

K. Porayska-Pomsta. 2016. AI as a Methodology for Supporting Educational Praxis and Teachers? Metacognition. *International Journal of Artificial Intelligence in Education* 26, 2 (2016), 679–700.

K. Porayska-Pomsta, K. Anderson, S. Bernardini, K. Guldberg, T. Smith, L. Kossivaki, S. Hodgings, and I. Lowe. 2013. Building and an intelligent, authorable serious game for autistic children and their carers.. In *Advances in Computer Entertainment.* Springer, Cham, 456–475.

K. Porayska-Pomsta, C. Frauenberger, H. Pain, G. Rajendran, T. Smith, R. Menzies, M. E. Foster, A. Alcorn, S. Wass, S. Bernardini, K. Avramides, W. Keay-Bright, J. Chen, A. Waller, K. Guldberg, J. Good, and O. Lemon. 2011. Developing technology for autism: an interdisciplinary approach. *Personal Ubiquitous Comput.* 16, 2 (Feb. 2011), 117–127.

B.M. Prizant, A.M. Wetherby, E. Rubin, and A.C. Laurent. 2003. The SCERTS model: A Transactional, Family-Centered Approach to Enhancing Communication and Socioemotional Ability in Children with Autism Spectrum Disorder. *Infants and Young Children* 16, 4 (2003), 296–316.

B.M. Prizant, A.M. Wetherby, E. Rubin, A.C. Laurent, and P.J. Rydell. 2006. *The SCERTS® Model: A Comprehensive Educational Approach for Children with Autism Spectrum Disorders.* Brookes.

M. M. Rahman, S. Ferdous, S. I. Ahmed, and A. Anwar. 2011. Speech development of autistic children by interactive computer games. *Interactive Technology and Smart Education* 8, 4 (2011), 208–223.

G. Rajendran. 2013. Virtual environments and autism: a developmental psychopathological approach. *Journal of Computer Assisted Learning* 29, 4 (2013), 334–347. https://doi.org/10.1111/jcal.12006

S. Ramdoss, R. Lang, A. Mulloy, J. Franco, M. O'Reilly, R. Didden, and G. Lancioni. 2011a. Use of computer-based interventions to teach communication skills to children with autism spectrum disorders: A systematic review. *Journal of Behavioral Education* 20, 1 (2011), 55–76.

S. Ramdoss, A. Mulloy, R. Lang, M. O'Reilly, J. Sigafoos, and et al. Lancioni, G. 2011b. Use of computer-based interventions to improve literacy skills in students with autism spectrum disorders: A systematic review. *Research in Autism Spectrum Disorders* 5, 4 (2011), 1306–1318.

P. A. Rao, D. C. Beidel, and M. J. Murray. 2008. Character-based interactive storytelling. *Journal of Autism and Developmental Disorders* 38, 2 (February 2008), 353–361.

V. S. Rao, V. Raman, and A. V.. Mysore. 2015. Issues related to obtaining intelligence quotient-matched controls in autism research. *Indian journal of psychological medicine* 37, 2 (2015), 149.

B. Reichow, F.R. Volkmar, and D.V. Chiccetti. 2008. Development of the evaluative method for evaluating and determining evidence-based practices in autism. *Journal of Autism and Developmental Disorders* 38 (2008), 1311–1319.

G. Reynhout and M. Carter. 2009. The use of Social Stories by teachers and their perceived efficacy. *Research in Autism Spectrum Disorders* 3, 1 (2009), 232–251.

S.J. Russell and P. Norvig. 2003. *Artificial Intelligence: A Modern Approach.* Prentice Hall. Second Edition.

N J. Salkind. 2010. *Encyclopedia of Research Design.* Sage Research Methods.

D. Scattone, D. H. Tingstrom, and S. M. Wilczynski. 2006. Increasing Appropriate Social Interactions of Children With Autism Spectrum Disorders Using Social StoriesTM. *Focus on Autism and Other Developmental Disabilities* 21, 4 (2006), 211–222.

B. Schuller, E. Marchi, S. Baron-Cohen, H. OReilly, P. Robinson, I. Davies, O. Golan, S. Friedenson, S. Tal, S. Newman, N. Meir, R. Shillo, A. Camurri, S. Piana, S. Blte, D. Lundqvist, B. Berggren, S. A., and N. Sullings. 2013. IASC-Inclusion: Interactive Emotion Games for Social Inclusion of Children with Autism Spectrum Conditions.. In *Proceedings 1st International Workshop on Intelligent Digital Games for Empowerment and Inclusion (IDGEI '13).*

C. A. Smith and R. S. Lazarus. 1990. Emotion and Adaptation. In *Handbook of Personality: Theory and Research*, I. Vlahavas and D. Vrakas (Eds.). L.A. Pervin (Ed.)., New York: Guilford, Chapter 23, 609–637.

A C. Stahmer, L. Schreimban, and AB. Cunningham. 2011. Toward a technology of treatment individualization for young children with autism spectrum disorders. *Brain Research* 1380 (2011), 229–239.

D. C. Strickland, D. McAllister, C. D. Coles, and S. Osborne. 2007. An Evolution of Virtual Reality Training Designs for Children With Autism and Fetal Alcohol Spectrum Disorders. *Topics in Language Disorders* 27, 3 (2007), 226–241.

J. Suh, A. Orinstein, M. Barton, C M. Chen, I M. Eigsti, N. Ramirez-Esparza, and D. Fein. 2016. Ratings of Broader Autism Phenotype and Personality Traits in Optimal Outcomes from Autism. *Journal of Autism and Developmental Disorders* 46, 11 (2016), 3505–3518.

A. Tartaro and J. Cassell. 2008. Playing with virtual peers: bootstrapping contingent discourse in children with autism.. In *Proceedings of the International Conference on the Learning Sciences (ICLA' 08).* ACM, 382–389.

Report US National Research Council. 2002. *Scientific Research in Education.* Technical Report.

S. Wass and K. Porayska-Pomsta. 2014. The uses of cognitive training technologies in the treatment of autism spectrum disorders. *Autism: International Journal of Research and Practice.* 18, 8 (2014), 851–871.

C. Whalen and L. Schreibman. 2003. Joint attention training for children with autism using behavior modification procedures. *Journal of Child psychology and psychiatry* 44, 3 (2003), 456–468.

M. Wooldridge and N. R. Jennings. 1995. Intelligent Agents: Theory and Practice. *Knowledge Engineering Review* 10, 2 (1995), 115–152.

B. Woolf. 2008. *Building Intelligent Tutoring Systems.* Morgan Kaufman.